Developing Ultrasound-Guided Intervention Technologies Enabled by Sensing Active Acoustic and Photoacoustic Point Sources

by

Alexis Cheng

A dissertation submitted to The Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy.

Baltimore, Maryland

July, 2017

Abstract

Image-guided therapy is a central part of modern medicine. By incorporating medical imaging into the planning, surgical, and evaluation process, image-guided therapy has helped surgeons perform less invasive and more precise procedures. Of the most commonly used medical imaging modalities, ultrasound imaging offers a unique combination of cost-effectiveness, safety, and mobility. Advanced ultrasound-guided interventional systems will often require calibration and tracking technologies to enable all of their capabilities. Many of these technologies rely on localizing point-based fiducials to accomplish their task.

In this thesis, I investigate how sensing and localizing active acoustic and photoacoustic point sources can have a substantial impact in intraoperative ultrasound. The goals of these methods are (1) to improve localization and visualization for point targets that are not easily distinguished under conventional ultrasound and (2) to track and register ultrasound sensors with the use of active point sources as non-physical fiducials or markers.

We applied these methods to three main research topics. The first is an ultrasound

ABSTRACT

calibration framework that utilizes an active acoustic source as the phantom to aid in in-plane segmentation as well as out-of-plane estimation. The second is an interventional photoacoustic surgical system that utilizes the photoacoustic effect to create markers for tracking ultrasound transducers. We demonstrate variations of this idea to track a wide range of ultrasound transducers (three-dimensional, two-dimensional, bi-planar). The third is a set of interventional tool tracking methods combining the use of acoustic elements embedded onto the tool with the use of photoacoustic markers.

Primary Supervisor: Emad M. Boctor

Secondary Supervisors: Russell H. Taylor, Gregory S. Chirikjian

Acknowledgments

This thesis is the culmination of the 6 years of my PhD studies in the MUSiiC lab. I want to thank Dr. Emad M. Boctor for his support in helping me grow as a researcher and an innovator throughout my years here. This thesis would not have been possible without him.

I want to thank Dr. Russell H. Taylor for being so generous with his advice and insights into being during our time together in both research and teaching.

I want to thank Dr. Gregory S. Chirikjian who has been an important mentor as we collaborated together on some of my first research projects.

This work would also not have been possible without support from my friends and colleagues in the MUSiiC lab and LCSR including: Dr. Xiaoyu Guo, Dr. Hyunjae Kang, Dr. M. Kendal Ackerman, Haichong K. Zhang, Younsu Kim, Qianli Ma, Dr. Jeeun Kang, Dr. Chloe Audigier, Dr. Nishikant Deshmukh, Dr. Fereshteh Aalamifar, Dr. Behnoosh Tavakoli, Dr. Muyinatu Lediju-Bell, Dr. Daniel Carnegie, Dr. Nathanael Kuo, Dr. Lei Chen, Dr. Seth Billings, Joshua Davis, Zerui Wang, Dr. Tsai Meng, Yuttana Itsarachaiyot, Singchun Lee, Long Qian, Li Chi, Xiang Xiang,

ACKNOWLEDGMENTS

Dr. Colin Lea, Dr. Sungmin Kim Nathan Cho, Taesoo Kim, Tutkun Sen, Rob Grupp.

Dedication

This thesis is dedicated to my family for their support throughout my academic career and personal life.

Contents

CONTENTS

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background

Image-guided therapy is a central part of modern medicine. By incorporating medical imaging into the planning, surgical, and evaluation process, image-guided therapy has helped surgeons perform less invasive and more precise procedures. Of the most commonly used medical imaging modalities, ultrasound imaging offers a unique combination of cost-effectiveness, safety, and mobility. Intraoperative ultrasound (IOUS) can provide real time two-dimensional and three-dimensional views into the patient's body with relatively simple hardware and minimal risks. There has also been an increasing variety of available IOUS transducers in the form of transcutaneous, laparoscopic, and catheter-based probes, increasing the possible procedures that may benefit from IOUS imaging. This has led to IOUS to be used in numer-

ous applications including laparoscopic surgery [2–4], guidance for biopsies [5], tumor ablation therapy [6], brachytherapy [7], catheterization [8], and robotic surgery [9].

Advanced IOUS-guided interventional systems will often require calibration and tracking technologies to enable all of their capabilities, such as real-time visualization of surgical tools inside the patient's body or fusion with preoperative medical information such as computed tomography (CT) or magnetic resonance imaging (MRI). This information is useful to surgeons as it helps them spatially relate what they see in preoperative medical images to the current surgical scene. In addition, it can also help surgeons identify the locations of intraoperative surgical tools such as needles and catheters, which may be difficult to visualize in conventional IOUS. These interventional guidance systems are becoming standard of care in some areas.

1.2 Prior Art

Calibration and tracking technologies remain a significant challenge, especially with the continued development of advanced ultrasound imaging modlities as well as image processing techniques. For example, photoacoustic (PA) imaging can provide certain optical properties of tissue, and can be used in conjunction with traditional ultrasound imaging. Thus far, photoacoustic imaging has only been rarely used for clinical applications [10] and one of the limiting factors is its lack of dedicated tool tracking technologies. A photoacoustic-based tool tracking method would be much

more easily integrated into photoacoustic imaging systems and may aid in the translation of such systems. As a more developed example, ultrasound elastography is used for monitoring ablation of liver tumors [11, 12], image-guided prostatectomies [13], and breast cancer radiotherapy targeting [14]. One of the more common approaches is to place tracking devices on the IOUS transducer as well as the other surgical instruments [15–17]. Many of these technologies may rely on localizing point-based fiducials to accomplish their task. One very common method to relate or register pre-operative and intraoperative imaging is through the use of fiducials. These methods require a set of rigid fiducials that can be observed in both imaging modalities, which can then be used to register the two modalities together. The most simple realization of these set of fiducials is a set of points. They are generally a set of passive points, whose appearance depends completely on the medical imaging modality used.

The tracking of tools such as ultrasound transducers or needles or catheters, will generally make use of external tracking sensors such as optical tracking or electromagnetic (EM) sensing [18] to provide real time spatial information of the tool relative to the patient. Optical tracking systems require line of sight, while EM-based systems are wired and subjet to EM field distortions, discouraging the use of metallic tools. In addition, the estimation of the tool tips is limited by tool shaft bending and the effects of angle estimation error if the sensors themselves are placed far away from the tip. Further, IOUS to camera or IOUS to tool tracking transformations necessarily require an indirect calculation based on a chain of spatial transformations, each with

errors that may propagate to the next. Finally some type of calibration, either ultrasound calibration or tool pivot calibration is required. The tools themselves are also often difficult to visualize within an IOUS image. Stoll et al. [19] attached passive markers on the surgical instrument such that its position and orientation could be determined from an ultrasound image. Rohling et al. explored image processing [20] and beamforming approaches [21] to enhance tool visibility.

1.3 Thesis Statement and Outline

To address these issues, I investigate how sensing and localizing active acoustic and photoacoustic point sources can have a substantial impact in intraoperative ultrasound. The goals of these methods are (1) to improve localization and visualization for point targets that are not easily distinguished under conventional ultrasound and (2) to track and register ultrasound sensors with the use of active point sources as non-physical fiducials or markers.

We apply these methods to three main impact areas. The first is an ultrasound calibration framework that utilizes an active acoustic source as the phantom to aid in in-plane segmentation as well as out-of-plane estimation. The second is an interventional photoacoustic surgical system that utilizes the photoacoustic effect to create markers for tracking ultrasound transducers. We demonstrate variations of this idea to track a wide range of ultrasound transducers (three-dimensional, two-dimensional,

bi-planar). The third is a set of interventional tool tracking methods combining the use of acoustic elements embedded onto the tool with the use of photoacoustic markers. These parts explore the use of active points in the context of ultrasound calibration, tracking of ultrasound transducer, and interventional tool tracking.

Chapter 3 in part I explores the use of active fiducials applied in an ultrasound calibration scenario and we examine methodologies and techniques that active point sources enable for this scenario. Chapters 4, 5, 6, and 7 in part II explores the use of photoacoustics to create non-physical markers to track a variety of ultrasound transducers. Different methods are used under different configurations. Chapters 8, 9, and 10 in part III demonstrate three methods for tool tracking based on active sensors or photoacoustics.

1.4 Contributions

1.4.1 Ultrasound Calibration with Active Phantoms

My major contributions to Part I are the following:

- First active phantom for ultrasound calibration with the ability to communicate with the ultrasound transducer [22]
- Image- and User- independent ultrasound calibration [22]

- Out-of-plane algorithm for point-based ultrasound calibration [23]

My minor contributions to Part I are the following:

- Z-wire and cross-sectional feature-based phantoms [24]
- Probabilistic and correspondence-free calibration methods [25–27]
- Automatic robotic ultrasound calibration [28]

1.4.2 Tracking Ultrasound Transducers in an Interventional Photoacoustic Surgical System (I-PASS)

My major contributions to Part II are the following:

- First Photoacoustic-based ultrasound transducer tracking systems [29–33]
- PA-based 3D Volumetric US Probe Tracking [29–31]
- PA-based 2D Convex US Probe Tracking [32]
- PA-based Bi-planar Transrectal US (TRUS) Probe Tracking [33]
- Out-of-plane PA marker localization [32, 33]

My minor contributions to Part II are the following:

- Sub-surface photoacoustic fiducials for intraoperative-preoperative registration [34]

1.4.3 Interventional tool tracking with photoacoustic sources and single element ultrasound receivers

My major contributions to Part III are the following:

- Interventional tool tracking with embedded piezoelectric elements [35]
- First photoacoustic-based catheter tracking method demonstrated in vivo [35]
- Virtual rigid body optical tracking with non-physical markers created with light [36]
- Needle tracking with fused optical and piezoelectric sensing

My minor contributions to Part III are the following:

- Robotic ultrasound servoing for catheter tracking [37]

Part I

Ultrasound Calibration with Active Phantoms

Image-guided surgery (IGS) systems are often used in modern surgical procedures to provide surgeons with additional information support and guidance leading to less trauma for the patient. Specific benefits to the patient can include cost reduction of the procedure, reduced morbidity rates, and shorter recovery times. In IGS systems, an intraoperative medical imaging modality is often used to provide a visualization of underlying tissue structures or anatomy that cannot be seen with the naked eye. In this part, we focus on the use of ultrasound (US) in IGS.

US imaging systems are widely integrated with tracking or robotic systems in IGS systems for tool tracking and image guidance. An US calibration is necessary before tracking information from tracking or robotic devices can be used in conjunction with US to perform more advanced forms of guidance. The first requirement for this process is that a tracked reference frame must be rigidly attached to the US transducer. This can be either an optical marker, an electromagnetic (EM) sensor, or a mechanically tracked attachment, such as a robotic arm. Although they have different error profiles, they can all be considered as external trackers in the context of US calibration, providing the tracked reference frames pose, its orientation and position, relative to itself. The US calibration process finds the rigid body transformation relating the tracked reference frame to the US image, allowing an US image to be positioned relative to the external tracker. The US image is now registered with anything else, such as other tools or devices, that is being tracked by this external tracker. Calibration enables more advanced uses of the US system, such as video and US image overlays or

the targeting of regions of interest with robotically-actuated tools. Relative to other imaging modalities such as magnetic resonance imaging (MRI) or computed tomography (CT), US suffers from poor image quality and limited field of views. These drawbacks often make it difficult to accurately localize features within the US image plane and automatically segment regions of interest within the images.

This part presents the work done related to ultrasound calibration. Chapter 2 discusses phantom and algorithmic development for passive ultrasound calibration phantoms. Chapter 3 explores the use of an active ultrasound phantom.

Chapter 2

AX=XB ultrasound calibration phantoms and algorithms

2.1 Introduction

To find the transformation between the tracked and image frames in the US calibration, a phantom with a known configuration is often required. There have been many different types of phantoms or models used for US calibration including wall [38], cross-wire (CW) [39], Z-fiducial [40], and $AX = XB$ [41] phantoms.

The US calibration phantom that we develop in this work can be used for both the Z-fiducial and $AX = XB$ calibration approaches. Z-fiducial phantoms are those where three wires are oriented in a plane to form a Z or N shape. Given a plane intersecting all three wires of multiple Z-fiducials, the unique pose of the plane can

be determined. There are two methods for using Z-fiducial phantoms in US calibration. In the first method, only a single image is required, but the phantom must be externally tracked in addition to the probe fiducial. In the second method, the phantom is not externally tracked, but a moderate number of images, from differing poses, are required. $AX = XB$ phantoms are those that allow the relative rigid body transformation between two images to be recovered based on each images content. Z-fiducial phantoms are a subset of $AX = XB$ phantoms and therefore also fit this requirement.

The $AX = XB$ problem is also referred to as the hand-eye calibration problem. As seen in figure 2.1, A^{ij} and B^{ij} are relative motions related by the rigid body transformation X. A tracker provides the homogeneous transformation B_i, representing the pose of the externally tracked marker or sensor. By using an $AX = XB$ phantom, the absolute pose of the image, A_i, relating each image to the phantoms coordinate system, can be computed based on the image content. The $AX = XB$ framework ultimately only requires relative motions and is therefore advantageous because it does not require the US transducer to be fixed at specific locations or the calibration phantom to be tracked by the external tracker. The calibration phantom must be fixed during the data collection process, but the US transducer can freely move, and data can be collected in real-time.

In this work we design and test a new calibration phantom. This calibration phantom improves on generic Z-fiducial phantoms by having a design that is more

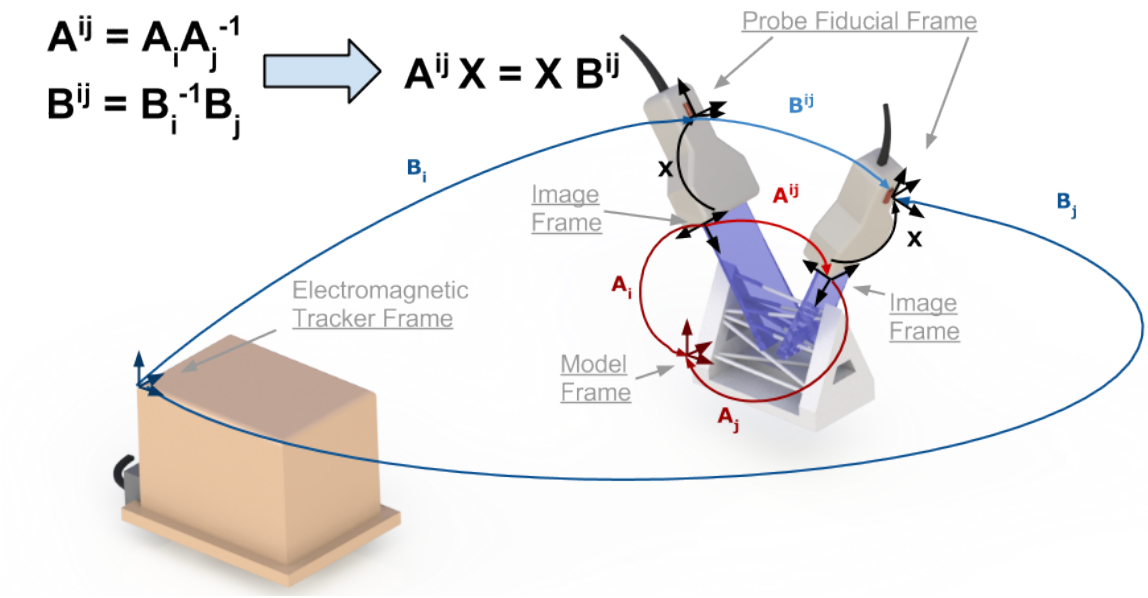


Figure 2.1: $AX = XB$ formulation with labeled coordinate frames defining absolute and relative pose transformations.

compatible with US physics. We describe the phantom design and the factors behind these decisions as well as the segmentation and registration software to support this calibration phantom. We present experimental results using this calibration phantom. We also demonstrate in simulation how Euclidean invariants from $AX = XB$ can be used to filter noisy data as well as solve for X with unknown correspondence.

2.2 Contributions

The main contributions of this chapter are the following:

- Development and validation of an ultrasound calibration phantom that compensates for limitations due to ultrasound physics

- Development and validation of Euclidean invariants under the $AX = XB$ framework for filtering and calibration with unknown correspondence

2.3 Acknowledgement

This work was conducted in close collaboration with my colleague, Dr. M. Kendal Ackerman. I led the phantom design and experiments, but also participated in the development and the analysis of the ultrasound calibration algorithms and data.

2.4 Technical Approach

This work consists of phantom design and an ultrasound calibration framework.

2.4.1 Phantom Design

The new calibration phantom that is presented in figure 2.2, where the phantoms computer aided design in figure 2.2A and the three-dimensional printed phantom in figure 2.2B are shown. There were several design factors that distinguish this phantom from previous phantoms. The first constraint was that the phantom be created without any need for additional assembly, adjustment, or treatment. Previous printed phantoms generally consist of a frame with holes, identifying the position and orientation of wires. The user would then wire off-the-shelf fishing line [6] into the

phantom. In contrast, our entire phantom, including the rods, is printed, allowing our phantom to be plug and play. The user can download the appropriate computer-aided design (CAD) file, print it from a three-dimensional printer, and be ready to perform US calibration. There are, however, limits to the specifications of a three-dimensional printer. It may be difficult to print thin wires as they can become brittle. For this reason, the printed wires are 2mm in diameter, thicker than off-the-shelf fishing line. Thick wires affect the US images in negative ways, making it more difficult to automatically segment. We will discuss this in the design of the US calibration software. The terms rods and wires will be used interchangeably in this manuscript.

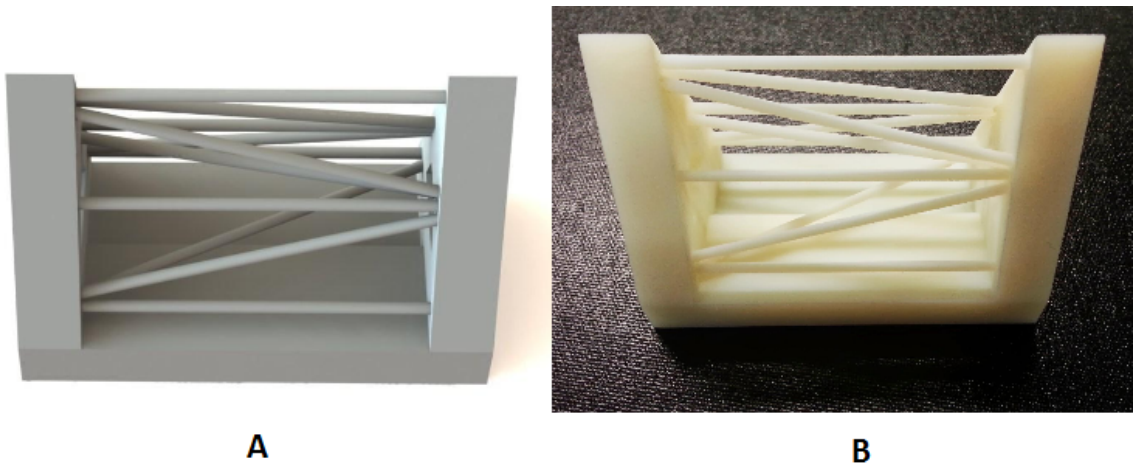


Figure 2.2: Image of calibration phantom A)CAD model B)Physical phantom.

Since this phantom is an extension of Z-fiducial phantoms, there were many considerations regarding the construction of the Z-fiducial. First, the length and angle of the Z-fiducial were chosen such that a submillimeter translation could be resolved.

In standard phantoms of this type, the Z-fiducials all lie on parallel planes. This geometry does not take advantage of US physics, as the axial dimension has higher resolution than the lateral dimension. In our phantom, a portion of the Z-fiducial change is reflected in the axial dimension, in varying orientations. The new geometry causes rods that are not perpendicular to the imaging plane to have a non-optimal acoustic response, leading to low rod intensities in the US image. To compensate for this, the Z-fiducials are oriented such that in cases where one Z-fiducial becomes difficult to visualize, another Z-fiducial will have a rod that becomes increasingly perpendicular to the image plane. A simulated situation is shown in figure 2.3A-B. The phantom also has many redundant rods, so that the subset of rods with the highest acoustic response, given the imaging orientation, can be chosen.

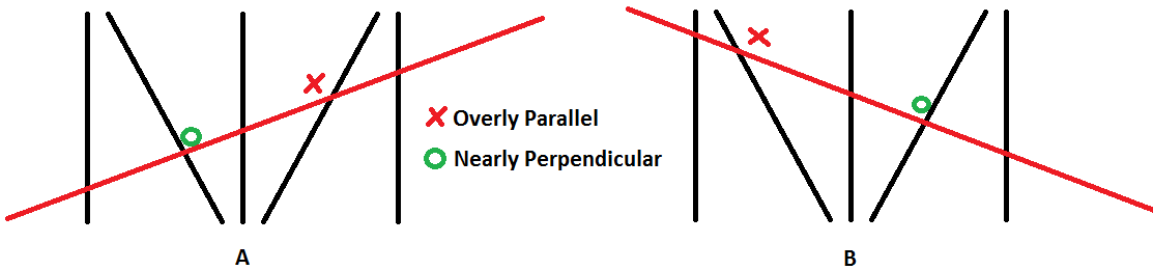


Figure 2.3: Orientation of Z fiducials A)Case 1 B)Case 2.

Another consideration was the shadowing effect of the structures within an US image. To avoid the case where many of the rods do not have a clear acoustic path to the transducer, the Z-fiducials are oriented in the shape of a triangle such that when one face of the triangle is experiencing severe shadowing effects, the other faces

will be unaffected. This phantom was also designed for use with probes of multiple lateral lengths. For example, with probes of shorter lateral lengths, one could use the six uppermost rods and perform the calibration. Finally, we desired the phantom to allow for a large range of motion of at least three cm for each translational degree of freedom and forty-five degrees for each rotational degree of freedom.

2.4.2 Ultrasound Calibration Framework

The calibration software has three essential components. Its general workflow can be found in figure 2.4A. The first is an automatic segmentation algorithm to detect the rods present in the US image. The second component computes the homogeneous transformation, A_i, relating each image with the phantom. The final component solves for the sensor calibration rigid body transformation, X.

The first component is an automatic segmentation algorithm. Its overall workflow can be seen in figure 2.4B. The first step of the segmentation algorithm applies an intensity threshold, chosen by Otsus method [42], to the image. This step allows the algorithm to get a first order estimate of which pixels should be considered the background and which pixels should be considered the signal. A connected regions algorithm is then used to cluster signal pixels together. Afterwards, a filter is applied where only regions containing a certain range of pixels are retained. These steps allow us to remove noise and extract the rods from the US image. After these steps, there will be multiple regions of pixels, indicating rod candidates. An example US image

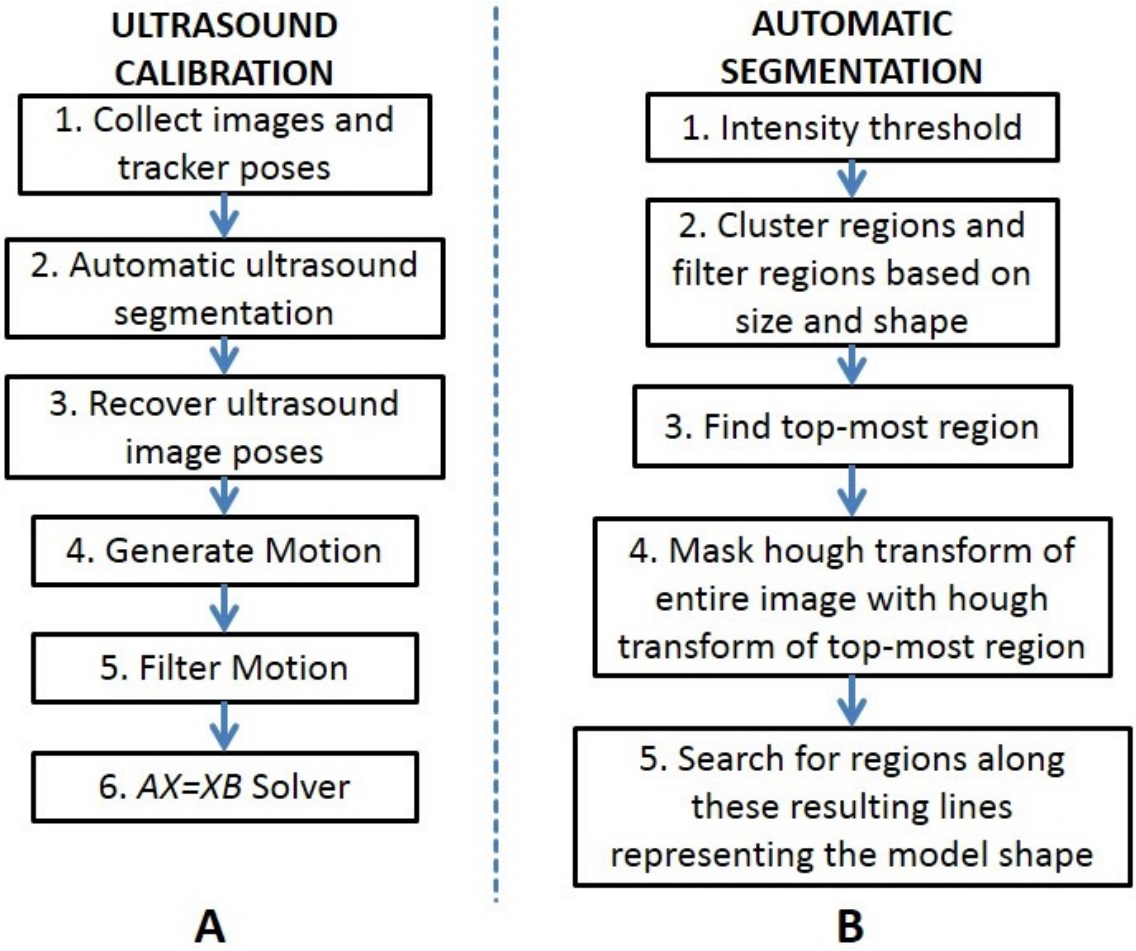


Figure 2.4: Workflows for A) the overall US calibration and B) the automatic segmentation algorithm

and the segmented rod candidates can be seen in figure 2.5. The next steps must find correspondence between these wire candidates and the model.

The region closest to the US transducer face is selected as the top rod in our model. This is a valid assumption, since in cases where the phantom is placed in a water bath, it is not practical to place the transducer such that the top rod is not closest to the US transducer face. Given our model, the remaining connected regions in the US image will exhibit a triangular shape. Thus, the standard Hough

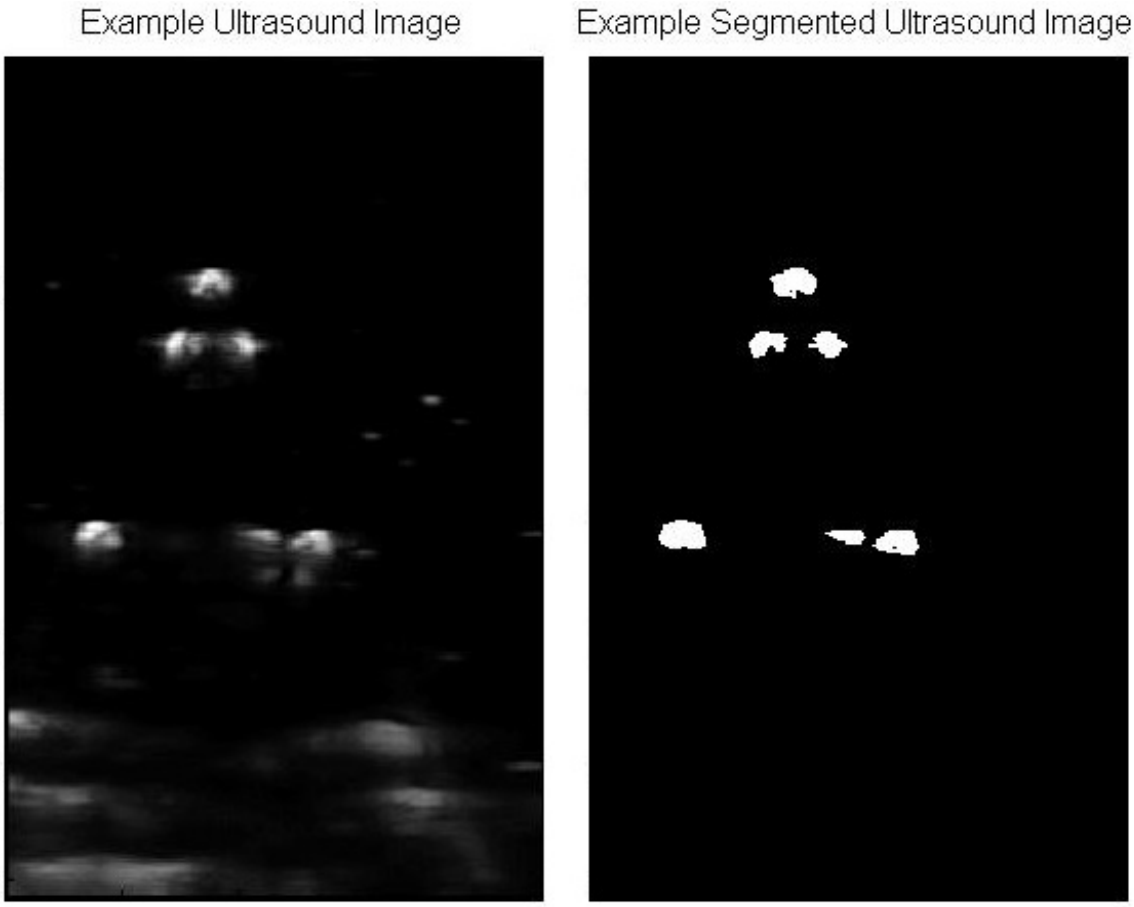


Figure 2.5: Example Ultrasound Image and Example Segmented Ultrasound Image

transform [43] can be applied to find the edges of this triangular pattern. We use our knowledge of the location of the top rod and the edges of the triangular pattern to establish correspondences between the triangular pattern and the model. With knowledge of these lines, we can search the original image to find rods that had low intensities. This step ensures that even if a rod was only faintly seen in the image, it will still be detected. Finally, for each region, we select the points lying closest to the transducer face from the centroid because the top of the rods are more accurately

represented in the US image.

We examined three registration methods to compute the homogeneous transformation, A_i. First, we can directly register these points to a discretized model using coherent point drift [44]. This algorithm does not require known correspondences and also allows for missing points. However, this algorithm is prone to local minima during its iterative procedure. Thus, it may be necessary to initialize the segmented points at different locations and hope a majority converge to the correct answer. The second method requires known correspondences and is based on the Z-fiducials. Given a plane that intersects all three rods of the Z-fiducial, the intersection of the plane and the diagonal rods can be localized in the phantom space. Thus, if a minimum of three Z-fiducials are visible in the US image, the transformation between the image space and the phantom space can be found. However, if certain rods are not visible, there may not be enough Z-fiducials to use this method. A third method uses the inter-point distances of the segmented points and fits this distance vector to the model. This method also requires correspondence. In the case when not all of the points exist in the image, a more complex fitting method, such as the Hungarian method [45], can be used. The fitting process can be accomplished through either an optimization procedure, which is also prone to local minima, or through a look-up table, which could require significant system memory. The preferred approach for this fitting process is to use a hybrid approach: a coarse look-up table is used to find an initial solution to the optimization procedure. The optimization procedure would

then be used to refine the solution towards its global minima.

2.4.3 $AX = XB$ Euclidean Invariants for Filtering

From screw theory it is known that any homogeneous transformation can be written as [46]

$$H = \begin{pmatrix} e^{\theta N} & (\mathbb{I}_3 - e^{\theta N})\mathbf{p} + d\,\mathbf{n} \\ \\ \mathbf{0}^T & 1 \end{pmatrix}$$

where $e^{\theta N}$ denotes the matrix exponential, $\mathbb{I}_n$ is the $n \times n$ identity matrix, and $\theta \in [0, \pi]$ is the angle of rotation.

$$N = \begin{pmatrix} 0 & -n_3 & n_2 \\ n_3 & 0 & -n_1 \\ -n_2 & n_1 & 0 \end{pmatrix}$$

where $\mathbf{n} = [\mathbf{n_1}, \mathbf{n_2}, \mathbf{n_3}]^{\mathbf{T}} \in \mathbb{R}^3$ is the unit vector describing the axis of rotation, which connects the origin and any point on the unit sphere, and $\mathbf{p} \cdot \mathbf{n} = 0$. Together, $\{\theta, d, \mathbf{n}, \mathbf{p}\}$ define the Plücker coordinates of the screw motion.

If we write $AX = XB$ as

$$A^i = XB^iX^{-1} \quad \text{where} \quad i \in \{1, 2\}, \tag{2.1}$$

then explicitly calculating and equating the matrix product gives two invariant relations,

$$\theta_{A^i} = \theta_{B^i} \quad d_{A^i} = d_{B^i} \tag{2.2}$$

where d_{A^i} and d_{B^i} are computed from A^i and B^i as in (2.1). Additionally, let

$$\mathbf{l}_{A^i}(t) = \mathbf{p}_{A^i} + t\mathbf{n}_{A^i} \text{ and } \mathbf{l}_{B^i}(t) = \mathbf{p}_{B^i} + t\mathbf{n}_{B^i}$$

be the directed screw axis lines of A_i and B_i in three-dimensional Euclidean space. If the lines are not parallel or anti-parallel, i.e., if $\mathbf{n}_{A_i} \neq \pm\mathbf{n}_{B_i}$, then the distance between the two lines is given by

$$\Delta(\mathbf{l}_{A^{i_1}}, \mathbf{l}_{A^{i_2}}) = \frac{|[\mathbf{n}_{A^{i_1}}, \mathbf{n}_{A^{i_2}}, \mathbf{p}_{A^{i_2}} - \mathbf{p}_{A^{i_1}}]|}{\|\mathbf{n}_{A^{i_1}} \times \mathbf{n}_{A^{i_2}}\|} \tag{2.3}$$

where for any $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$, the triple product is $[\mathbf{a}, \mathbf{b}, \mathbf{c}] \doteq \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$. In the current context we can think of $i_1 = 1$ and $i_2 = 2$ but in later discussion i_1 and i_2 can represent more general values.

If in addition, $\Delta(\mathbf{l}_{A^{i_1}}, \mathbf{l}_{A^{i_2}}) \neq 0$, i.e., if the lines are skew, then the angle $\phi(\mathbf{l}_{A^{i_1}}, \mathbf{l}_{A^{i_2}}) \in [0, 2\pi)$ is uniquely specified by

$$\cos\phi(\mathbf{l}_{A^{i_1}}, \mathbf{l}_{A^{i_2}}) = \mathbf{n}_{A^{i_1}} \cdot \mathbf{n}_{A^{i_2}} \tag{2.4}$$

$$\sin\phi(\mathbf{l}_{A^{i_1}}, \mathbf{l}_{A^{i_2}}) = \Delta(\mathbf{l}_{A^{i_1}}, \mathbf{l}_{A^{i_2}})^{-1}[\mathbf{n}_{A^{i_1}}, \mathbf{n}_{A^{i_2}}, \mathbf{p}_{A^{i_2}} - \mathbf{p}_{A^{i_1}}].$$

Therefore, if $\theta_{A^{i_1}}, \theta_{A^{i_2}} \in (0, \pi)$ and $\phi(\mathbf{l}_{A^{i_1}}, \mathbf{l}_{A^{i_2}}) \notin \{0, \pi\}$, then a unique solution of $AX = XB$ exists if and only if the following four conditions hold:

1. $\theta_{A^{i_1}} = \theta_{B^{i_1}}$ and $\theta_{A^{i_2}} = \theta_{B^{i_2}}$;

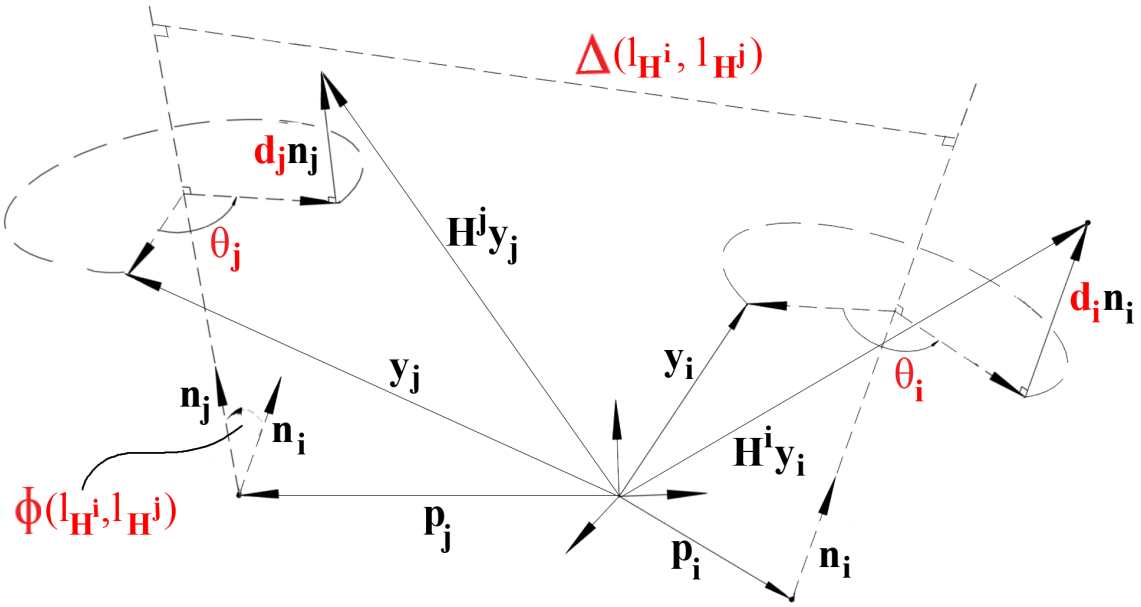


Figure 2.6: Two arbitrary rigid-body motions, H^i which is shown acting on y_i and H^j which is shown acting on y_j, their Plücker coordinates and the parameters of the four described conditions (in red) [1]

2. $d_{A^{i_1}} = d_{B^{i_1}}$ and $d_{A^{i_2}} = d_{B^{i_2}}$;

3. $\phi(\mathbf{l}_{A^{i_1}}, \mathbf{l}_{A^{i_2}}) = \phi(\mathbf{l}_{B^{i_1}}, \mathbf{l}_{B^{i_2}})$;

4. $\Delta(\mathbf{l}_{A^{i_1}}, \mathbf{l}_{A^{i_2}}) = \Delta(\mathbf{l}_{B^{i_1}}, \mathbf{l}_{B^{i_2}})$.

If these do not hold, then a solution will not be possible [47]. Figure 2.6 illustrates the Plücker coordinates, and the parameters of the above four conditions for two arbitrary rigid-body motions [1]

Before we can solve the $AX = XB$ problem, we must obtain the relative A and B transformations from their respective absolute poses. We can either take every consecutive pair where $A^i = A_i^{-1} A_{i+1}$ or every pair combination where $A^{ij} =$

$A_i^{-1}A_j$. In general, every pair will have better results than consecutive pairs at the expense of computation time, as there will be significantly more input data to the chosen $AX = XB$ solver. In addition, we applied a filtering step using the Euclidean invariants listed above of the homogeneous transformations in $AX = XB$, filtering small motions more affected by noise. We also filtered motions where the invariants were not sufficiently close. With relative A and B pairs, we can solve the $AX = XB$ problem using various solvers such as the dual quaternion method [48] or the kronecker product solver [49]. For this chapter, we used the kronecker product solver.

2.4.4 Recovering Unknown Correspondence

Suppose that data streams of sensor measurements $\mathbf{A} = \{A^i\}$ and $\mathbf{B} = \{B^j\}$ are presented and there are both significant unknown temporal shifts between these two sets, and gaps within each one. The number of points in these sets are $|\mathbf{A}| = m$ and $|\mathbf{B}| = n$.

Here we present an approach to recovering X and establishing a correspondence between the subsets $\mathbf{A}' \subset \mathbf{A}$ and $\mathbf{B}' \subset \mathbf{B}$ that do correspond where $|\mathbf{A}'| = |\mathbf{B}'| = p \leq \min(m, n)$. For such data, we find the correspondence, which is a permutation on p letters, $\pi \in \Pi_p$, such that $A^iX = XB^{\pi(i)}$ for $i = 1, ..., p$ where $A^i \in \mathbf{A}'$ and $B^{\pi(i)} \in \mathbf{B}'$.

We accomplish this using the invariants of the Special Euclidean group, $SE(3)$, under conjugation. The procedure is as follows. Compute (θ^{A^i}, d^{A^i}) for each $A^i \in \mathbf{A}$

and (θ_{B^j}, d_{B^j}) for each $B^j \in \mathbf{B}$. Next, form a 2D grid on the θ-d plane that ranges from $\min_{i,j}(\theta_{A^i}, \theta_{B^j})$ to $\max_{i,j}(\theta_{A^i}, \theta_{B^j})$ and $\min_{i,j}(d_{A^i}, d_{B^j})$ to $\max_{i,j}(d_{A^i}, d_{B^j})$. This grid will give r rectangles, e.g., if it is a 10×10 grid, then $r = 100$. Assuming that no data falls exactly on a grid line, this will partition $\mathbf{A}$ and $\mathbf{B}$ into r disjoint subsets: $\{\mathbf{A}_1, \mathbf{A}_2, ..., \mathbf{A}_r\}$ and $\{\mathbf{B}_1, \mathbf{B}_2, ..., \mathbf{B}_r\}$ where

$$\mathbf{A}_{i_1} \cap \mathbf{A}_{i_2} = \emptyset \quad \text{and} \quad \bigcup_{i=1}^{r} \mathbf{A}_i = \mathbf{A},$$

and similarly for $\mathbf{B}$.

The reason for doing this is that all candidate A's and B's that can potentially match will be in corresponding partitions $\mathbf{A}_i$ and $\mathbf{B}_i$, since having the same value of θ and d is a necessary condition for a solution to $AX = XB$ to exist. Constructing the grid with finite resolution allows for the possibility of some measurement error in A's and B's.

Let $|\mathbf{A}_i| = m_i$ and $|\mathbf{B}_j| = n_j$. Then

$$\sum_{i=1}^{r} m_i = m \quad \text{and} \quad \sum_{i=1}^{r} n_j = n.$$

Pick two bins for which all of the numbers in the pairs (m_{i_1}, n_{j_1}) and (m_{i_2}, n_{j_2}) are small, but greater than 2, to allow for the fact that measurement error may result in incorrect binning, and also that the angle $\phi(\mathbf{l}_{A^{i_1}}, \mathbf{l}_{A^{i_2}})$ might not always be in the range $(0, \pi)$. We interrogate all $m_{i_1} \times n_{j_1} \times m_{i_2} \times n_{j_2}$ possibilities as candidates. The further necessary conditions for the existence of a solution are $\phi(\mathbf{l}_{A^{i_1}}, \mathbf{l}_{A^{i_2}}) = \phi(\mathbf{l}_{B^{j_1}}, \mathbf{l}_{B^{j_2}})$ and $\Delta(\mathbf{l}_{A^{i_1}}, \mathbf{l}_{A^{i_2}}) = \Delta(\mathbf{l}_{B^{j_1}}, \mathbf{l}_{B^{j_2}})$.

2.5 Methods

2.5.1 Experiments

We use a Sonix RP (Ultrasonix Corp.) US machine, a L12-5 US transducer, an EM tracking system (Ascension Technology Corp.). The EM tracking system provides B_i, which is synchronized with the US images using the MUSiiC Toolkit [15]. This apparatus allows us to collect a continuous real-time stream of data. We performed the experiment in a water bath. In this experiment, we collect ultrasound images continuously. This allows us to collect synchronized pairs of electromagnetic tracking data and ultrasound images.

2.5.2 Simulation

A simulation was setup to validate the binning method for solving for $AX = XB$ when there is unknown correspondence. Noise-free $AX = XB$ data is generated with varying amounts of shift and gaps. Shifts correspond to a percentage where one data stream leads the other. Gaps correspond to data missing from a particular data stream.

2.6 Results

2.6.1 Experiments

Figure 2.7 shows the reconstructed phantom after calibration. If we denote the segmented phantom points as p_k, then the phantom model can be reconstructed by plotting $B_i X p_k$ for all combinations of i and k.

To validate this phantom, we fitted the phantom model to the reconstructed points and computed a normalized metric. For each point, we found its closest point on the model and computed the sum squared difference between them. We tested combinations of A^i or A^{ij} and with and without filtering. A set of test data not used in the computation of X is used to compute the results seen in table 2.1.

Motion Generation	Filtering	Error Metric (mm)
A^i	Without	2.75 ± 1.67
A^i	With	1.74 ± 1.00
A^{ij}	Without	2.36 ± 1.49
A^{ij}	With	1.56 ± 1.02

Table 2.1: Normalized error metric for different combinations of motion generation and filtering

2.6.2 Simulation

Figure 2.8 shows the success of this algorithm with different amounts of shift and gaps in the data. The percentages of shifts correspond with the percentage of data that does not overlap between the two data streams. The percentage of gaps correspond to the percentage of data that is missing from either of the two shifted data streams. Clearly the algorithm is highly robust to unknown and missing correspondences of all kinds.

2.7 Discussion

We can see from the results shown in figure 2.7 and table 2.1 that this phantom is a feasible US calibration phantom. A closer look at table 2.1 shows that the third and fourth rows respectively have lower errors than the first and second rows. This reinforces our initial hypothesis that motion generation from all pairs generally provides better results than motion generation from consecutive pairs. However, more experiments may be necessary to reach conclusive results. It can also be seen in table 2.1 that the second and fourth rows respectively have lower errors than the first and third rows. This corresponds with the notion that filtering using the Euclidean invariants [26] decreases the error.

There is certainly still room for improvement. US images are difficult to segment, and there are cases where the US image quality will cause our automatic segmentation

algorithm to fail. Since we have a continuous stream of data, there will eventually be enough data that can be automatically segmented. However, since the software is not currently real-time, it is impossible to know during data collection if there is enough data or not. Therefore, we aim to implement the software in real-time such that users could receive real-time feedback as they are collecting data.

2.8 Conclusion

In this chapter, we presented the development of an US calibration phantom that takes advantage of US physics and can be easily printed without user modifications. In addition, we describe a segmentation algorithm that leverages the shape and structure of our phantom. Finally, we show results of using this phantom in various combinations of motion generation and filtering methods. Future work will compare this phantom with other US calibration phantoms, develop a real-time US calibration framework using this phantom, and use this phantom as a test bed for other novel US calibration methods.

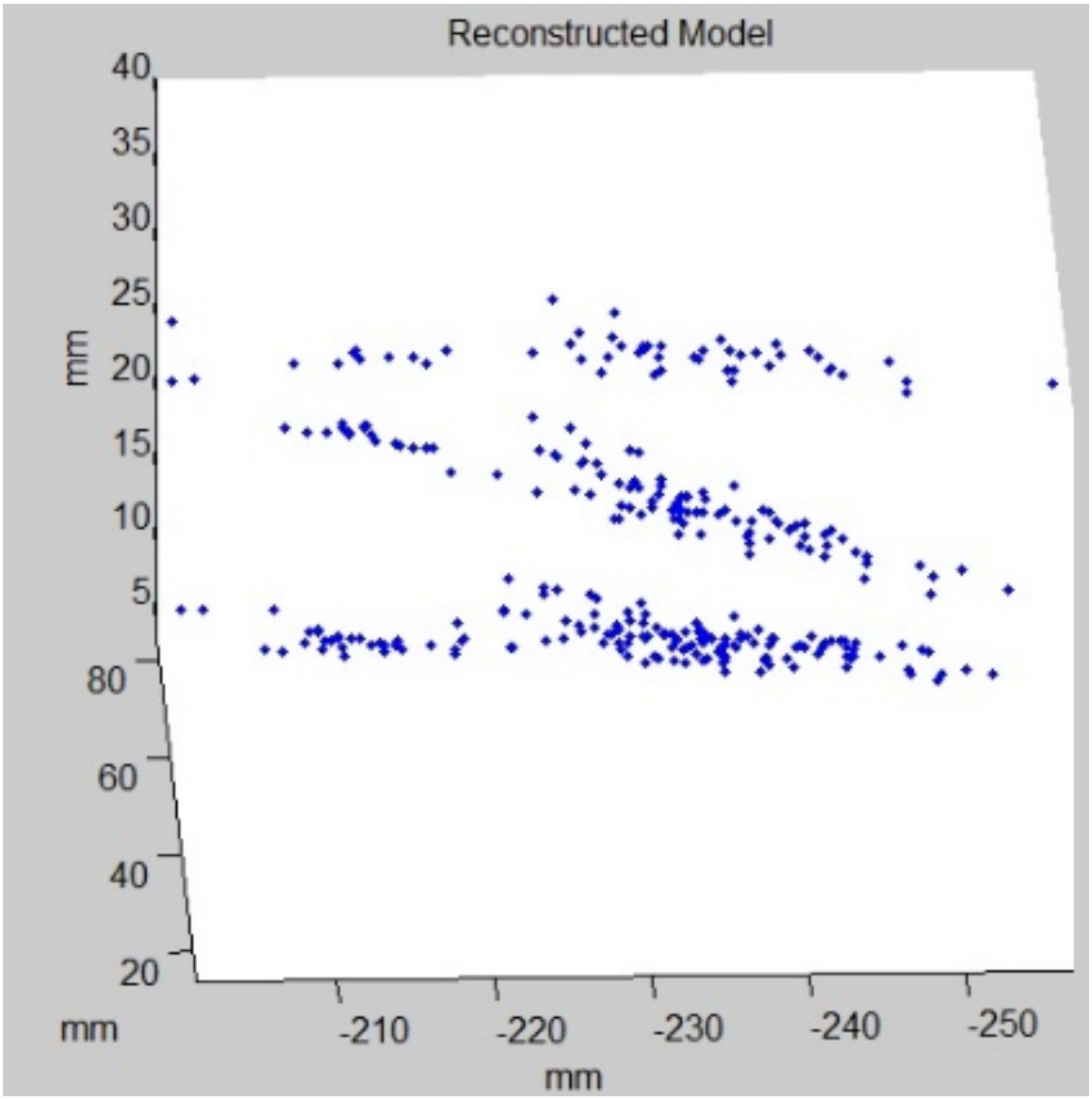


Figure 2.7: Example Ultrasound Image and Example Segmented Ultrasound Image

SHIFT \ GAPS	0%	10%	20%	30%	40%	50%	60%	70%	80%
0%	*100*	*100*	*100*	*100*	*100*	*98*	*58*	*13*	*2*
10%	*100*	*100*	*100*	*100*	*100*	*94*	*48*	*11*	*0*
20%	*100*	*100*	*100*	*100*	*99*	*80*	*40*	*4*	*0*
30%	*100*	*100*	*98*	*100*	*92*	*64*	*18*	*1*	*1*
40%	*100*	*100*	*100*	*99*	*91*	*53*	*20*	*1*	*0*
50%	*100*	*100*	*100*	*99*	*78*	*46*	*9*	*1*	*1*
60%	*100*	*100*	*99*	*96*	*62*	*23*	*1*	*0*	*0*
70%	*100*	*99*	*96*	*74*	*22*	*0*	*0*	*0*	*0*
80%	*100*	*99*	*65*	*4*	*0*	*0*	*0*	*0*	*0*

*Italicized data is percentage of trials that correctly solve for X, given the indicated shift and gaps (100 total trials).

Figure 2.8: Binning method Success Rate for Varying Amounts of Shift and Number of Small Gaps

Chapter 3

Active phantoms: a new paradigm for ultrasound calibration using phantom feedback

3.1 Introduction

In the previous chapter, I presented ultrasound calibration phantoms and methods related to $AX = XB$ phantoms. While ultrasound calibration under the AXXB framework allowed us to apply special methods for resolving correspondence and temporal synchronization, they also depended greatly on accurate segmentation of the phantom features within an ultrasound image. Segmentation error in each of the observed features can result in a compounded error in the computed image pose with

respect to the phantom. As such, there may be advantages with a BXp framework where only a single feature is required in each image.

In this form of US calibration, p is the fiducial point in image coordinates, B is the transformation measured by the external tracking system, and X is the unknown desired homogeneous transformation. The limitation of this method is that each recorded pose, B, must result in the fiducial point being seen by the US transducer. Given this requirement, each pair of B_i and p_i will represent the same physical point. This relationship can be described as shown in equation 3.1.

$$\forall i, j : B_i X p_i = B_j X p_j \tag{3.1}$$

In our BXp US calibration scenario shown in figure 3.1, US images of a static fiducial point are accumulated over various poses measured by a robotic arm. One then uses these poses and the segmented points in the US images to reconstruct a single point in the external trackers coordinate system. A limitation that prevents one from getting good calibration accuracy using BXp-based methods is the US transmission beam thickness. As we mentioned previously, this method requires an accurate segmentation of the static fiducial with respect to the US image. Accurate segmentation of the fiducial can be a difficult problem, especially in the elevational dimension due to the thickness of the US transmission beam. Depending on the depth and other imaging parameters, this beam can have a thickness ranging from several millimeters to centimeters, making it challenging to distinguish whether an object

in the B-mode image is intersecting the imaging mid-plane. This problem is further exacerbated by the fact that most calibration solvers do not account for out-of-plane errors. Since the localization and segmentation completely rely on the US image in conventional calibration phantoms, the elevation axis positioning uncertainty coupled with the relatively low quality of US image result in a reconstruction precision that can easily be worse than a few millimeters. Moreover, this is a user dependent procedure as the operator's experience in evaluating when the fiducial is in the mid-plane greatly affects the calibration accuracy.

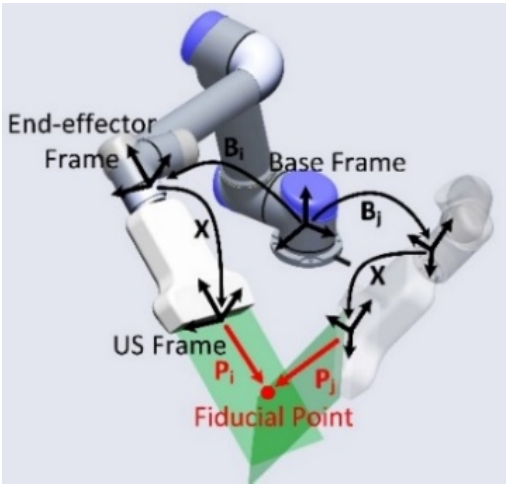


Figure 3.1: The coordinate transformations in the BXp calibration problem using a robotic arm as the tracking system. The US probe is attached to the robotic arm, and B_i is measured from the robot encoders.

Guo et al. [50] demonstrated the active ultrasound pattern injection system (AUSPIS). AUSPIS is an interventional tool tracking and guiding technique, one aspect of which solves the mid-plane error problem. In AUSPIS, an active echo (AE) element, which is a piezoelectric element (PZT) that acts as a single element US transducer, is

integrated with the target object that needs to be tracked in US images. An accompanying electrical system is used to control the AE element and receive data from it. During the US image acquisition process, probe elements fire sequentially to scan the entire field of view (FOV). If the AE element is in the FOV, it will be able to sense each of these pulses independently, if the signal intensity received at the AE element is sufficiently high. By transmitting an US pulse directly after each sensed event, AUSPIS can be used to improve the tool visualization. The US pulse transmitted by the AE element will be superimposed on the original reflected wave, resulting in an enhanced echo pulse with a much higher amplitude, configurable frequency and wider emission angle. This enhanced echo pulse can be seen in the US B-mode image directly at the location of the AE element. Another function of AUSPIS is to localize the AE element in the US mid-plane by measuring the local US signal intensity. This concept is shown in figure 3.2. The acoustic pressure is strongest in the mid-plane, thus the signal observed by the AE element can be used to determine when the AE element is in the US mid-plane. We use this feature to develop a search procedure that will be described in section 3.4.1. Since the AE element is a point that can be localized in an US image accurately, especially along the elevation axis, it is possible to use it in the same way as the CW point for US calibration.

An active US calibration phantom has several advantages when compared to passive US calibration phantoms. Besides image enhancement for segmenting within the image as well as accurate positioning in the elevational axis as described above, an

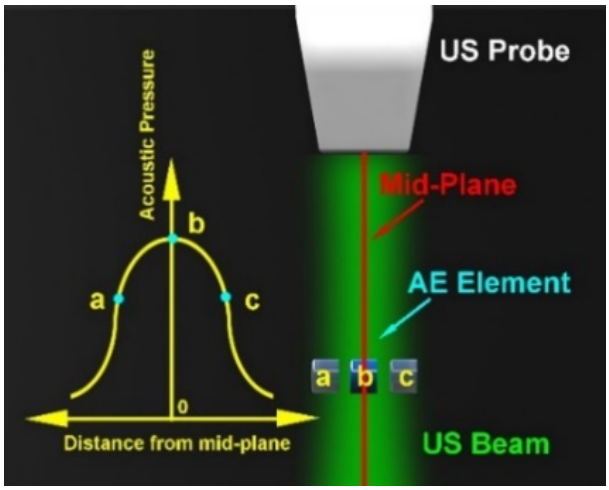


Figure 3.2: The main concept in using an AE element for mid-plane localization. At position b, AE element receives the maximum ultrasound signal amplitude. At a and c, although the element is still shown US image, the received signal amplitude is lower compared to b.

active phantom can also provide feedback to the user. This feedback can be used to develop automatic segmentation methods that can eliminate image dependency. If a controllable actuator, such as a robot arm, is used to position the US transducer, this feedback can also allow for control strategies that can eliminate user dependency. Aalamifar et al. [28] demonstrated an automatic calibration strategy using some of the hardware and methods described in this work.

In this chapter, we present the use of AUSPIS and its extension for placing targets within the US mid-plane and for automatically segmenting these targets within the image. How these methods contribute to US calibration with out-of-plane estimation will be described in section 3.4. Experimental procedures and data analysis protocols will be presented in section 3.5. The simulation and experimental results and subsequent discussion will be shown in sections 3.6.

3.2 Contributions

The main contributions of this chapter are the following:

- Development of an ultrasound calibration phantom that can actively transmit and receive acoustic signals with the ultrasound transducer
- Experimentation to show user-independence and image-independence using an active ultrasound calibration phantom
- Demonstration of an ultrasound calibration algorithm to compensate for out-of-plane uncertainty

3.3 Acknowledgement

This work was conducted in close collaboration with my colleague, Dr. Xiaoyu Guo. I led the experiments and the development and the analysis of the ultrasound calibration algorithms and data.

3.4 Technical Approach

The following sub-sections present methods for utilizing an active phantom for ultrasound calibration. The first two sections relate to in-plane ultrasound calibration whereas the two subsequent sections relate to ultrasound calibration with out-of-plane correction.

3.4.1 Mid-plane active point placement

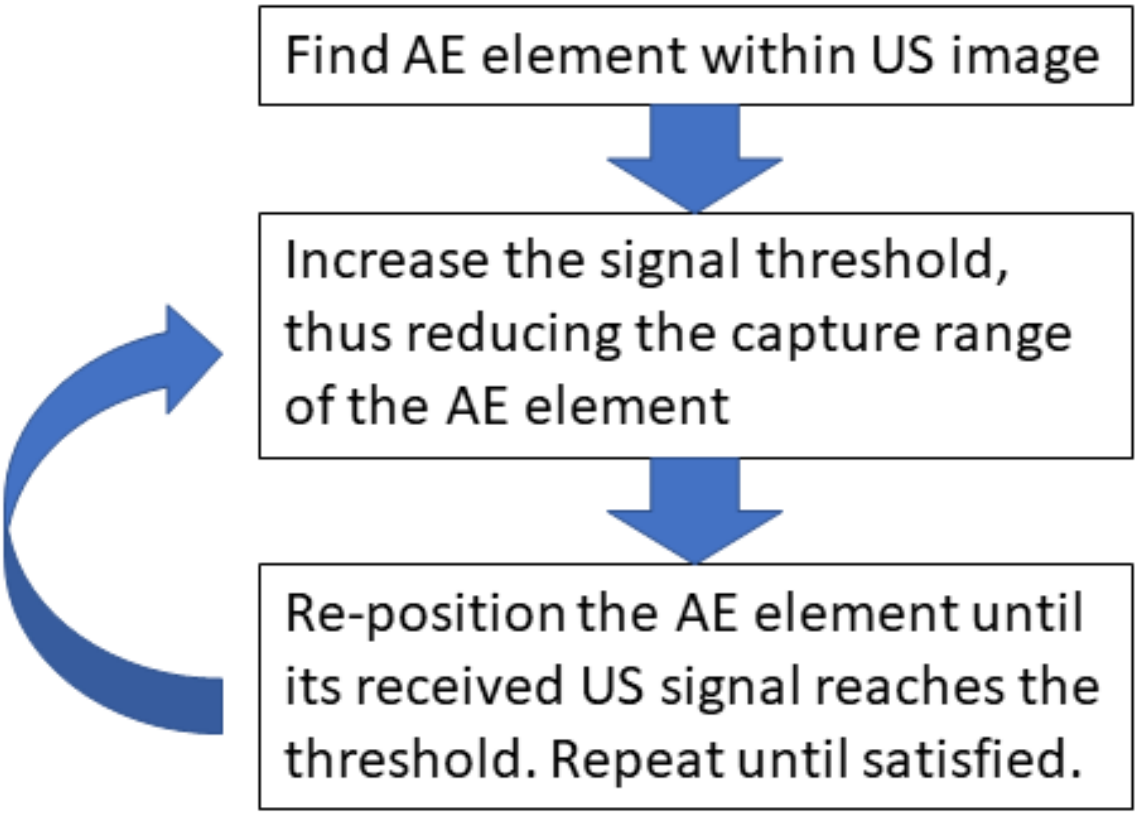


Figure 3.3: A flow chart demonstrating the procedure for placing the active point within the US mid-plane

One of the main issues with traditional US calibration is that the point target is assumed to be perfectly within the imaging mid-plane. This leads to the point targets seen in the B-mode image to have an elevational component of 0. Given that the AE element can be accurately placed in the US image mid-plane using AUSPIS, a natural use for the AE element is as the point target in US calibration to overcome this issue. We will describe two of many practical procedures for localizing the AE element within the imaging mid-plane. The general idea of this procedure is presented in the

three steps shown in figure 3.3. The first step is coarse localization, which follows conventional practice and is simply moving the US probe until the AE element can be seen. The second step is fine localization, which uses the signal amplitude received by the AE element as feedback. The AE element will respond to the US transmission only if the received signal amplitude has exceeded the pre-selected AE response intensity threshold. Initially, we use a low AE response intensity threshold, meaning that the AE element will respond within a larger distance from the imaging mid-plane than if the AE response intensity threshold was high. The third step is then to iteratively adjust this AE response intensity threshold and re-position the AE element, such that it gradually reaches the imaging mid-plane and any small motion results in the AE element not responding to the US transmission. Another approach to finding the imaging mid-plane is to use the virtual pattern injection technique present within AUSPIS. The basic idea is that AUSPIS can be used to inject an arbitrary pattern into the US image. This means that if we vary the injected pattern based on the received signal amplitude by the AE element, like cellular reception, this pattern can be used as a visual cue by the user. Figure 3.4 demonstrates this pattern injection approach.

More accurate and user independent positioning accuracy can be achieved along the elevation axis using AUSPIS for mid-plane positioning of the AE element. The user no longer must solely rely on their experience in analyzing US images as they will have additional feedback from AUSPIS. In other words, the point segmented from the

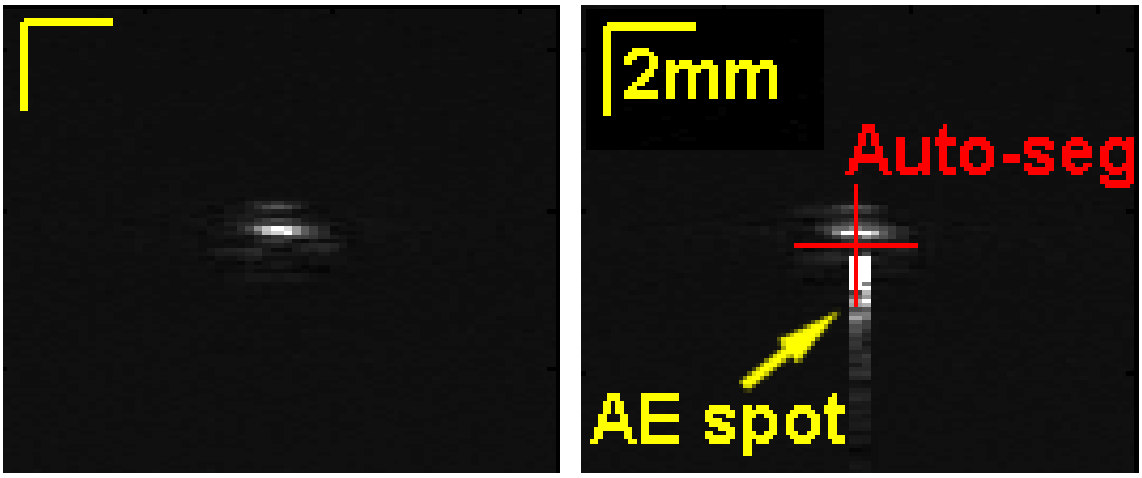


Figure 3.4: An ultrasound image demonstrating the pattern injection approach for mid-plane placement. The only change in the setup between the two ultrasound images shown is the operation of the AUSPIS feedback. The red cross also indicates the automatically segmented point.

US image and used in the calibration procedure now has an elevational component that is closer to the assumed 0 than traditional point positioning. Thus, one can hypothesize that a better and more consistent calibration precision can be obtained using AUSPIS as the static fiducial point in a BXp US calibration procedure.

3.4.2 Automatic active point segmentation

Point segmentation in US images depends heavily on the expertise of the user and the quality of the image. This means that the segmented point location may differ depending on what image settings were used and who determines the specific point location even if the physical relationship between the point and the US image are unchanged. For these reasons, it is beneficial to have an automatic point segmentation

method that is not based on image intensities found in the US images themselves. An active US calibration phantom is able to provide both US and electrical feedback when the US probe is aligned with the target point. This feedback can be used for automatic point segmentation. The US system outputs two triggers corresponding with the timing of the RF frame and each RF line respectively. AUSPIS receives these two triggers and compares them with the timing of the signal received by the AE element. Theoretically, this results in a time of flight (TOF) between the AE element and each of US transducer elements. The shortest of these TOFs will correspond to the US transducer element that is closest to the AE element. The distance from this US element to the AE element can be determined by multiplying this shortest TOF with the speed of sound (SOS) in the medium, as shown in equation 3.2. For a linear transducer, this distance can be directly used as the axial dimension of our segmented point. A Cartesian coordinate conversion would be necessary if the transducer is curvilinear. The index, i, of the US element with the shortest distance is also known by using the RF frame and line triggers. For a linear transducer, the lateral dimension can then be computed by multiplying i with the probe pitch, L_{pitch}, as shown in equation 3.3. By using this method, the point segmentation becomes entirely image and user-independent. An example can be seen in figure 3.4.

$$z = TOF_{shortest} * SoS \tag{3.2}$$

$$x = L_{pitch} * i_{shortest} \tag{3.3}$$

3.4.3 Out-of-plane estimation

While we have proposed a method to more accurately place the active point within the US image plane, there still remains some level of uncertainty. Conventional calibration solvers do not account for any elevational uncertainty that may exist in point localization. We propose a method to estimate and compensate for out-of-plane deviations. While the presented formulation is based on any point-based calibration data, we will also include an extension exclusive to AUSPIS or another apparatus capable of providing some semblance of pre-beamformed channel data.

As shown in figure 3.5, d_i represents the distance from the point to the US transducer for each element i. e_i represents the position of transducer element i in its local coordinate frame, which is computed by multiplying the element index with the transducer pitch. p represents the 3D position of the active point in the transducer coordinate frame. In the scenario where the lateral dimension of the point has been computed, through manual or automatic segmentation, only a single d_i corresponding to the shortest distance is known, again through manual or automatic segmentation. p must lie on the circle indicated on figure 3.5 to satisfy both the distance and closest element constraints. The general formulation is shown in equation 3.4 and any non-linear optimization method can be used to solve it. Figure 3.5 and equation 3.4

can be applied to both linear and curvilinear transducers, but there is a significant difference in the minimization result.

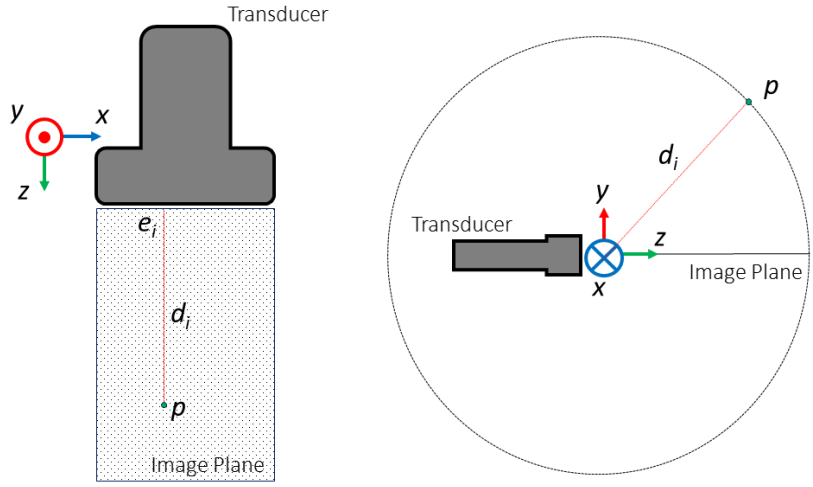


Figure 3.5: The concept for performing out-of-plane estimation. Given the lateral coordinate and the distance between the point and the transducer element closest to it, the point must exist on a circle within the axial-elevational plane.

$$d_i^2 = (||e_i - p||)^2 d_i^2 = (e_{ix} - p_x)^2 + (e_{iy} - p_y)^2 + (e_{iz} - p_z)^2 \tag{3.4}$$

$$d_i^2 = (e_{ix} - p_x)^2 + p_y^2 + p_z^2 \tag{3.5}$$

Equation 3.5 represents the set of distance equations between the active point and each of the transducer elements when using a linear transducer. Since the transducer elements are on a line, there is no axial or elevational component. We can easily see that there is one degree of freedom to satisfy this equation since any point rotated about the line defined by the transducer elements will have the same set of distances. The shortest distance between the transducer elements and the point results at $e_{ix} =$

p_x. In this case, we can see that equation 3.5 becomes the equation of a circle in the elevational-axial plane at $e_{ix} = p_x$. From this geometry, we can also see that any d_i not corresponding to $e_{ix} = p_x$ will not give any additional information to this problem. Thus, if given the lateral and axial position of any point in the image plane, we can also approximate a circle in the axial-elevational plane on which the point lies without using the pre-beamformed channel data that indicates the points distance to every element.

$$d_i^2 = (e_{ix} - p_x)^2 + p_y^2 + (e_{iz} - p_z)^2 \tag{3.6}$$

Equation 3.6 represents the set of distance equations between the active point and each of the transducer elements when using a curvilinear transducer. We can see that the main difference between this equation and equation 3.5 is the presence of an axial component in the transducer elements. This has a significant effect on the outcome as the degree of freedom mentioned above no longer exists. However, since the geometry is symmetrical on either side of the imaging plane, one can only solve for the elevational component up to a positive or negative sign. The additional information from the transducer geometry of a curvilinear transducer will restrict the location of the point to a single point on the circle, either in front of or behind the US image plane. It should be noted that this method of out-of-plane point estimation using a curvilinear transducer does require pre-beamformed channel data with the points distance to some or all of the US transducers elements.

3.4.4 Out-of-plane ultrasound calibration algorithm

The calibration algorithm will be described in the context of using a linear transducer, but it can also be applied in the curvilinear case. As previously mentioned, there is a circle for each particular image where equation 3.5 is satisfied. A standard ultrasound calibration algorithm looks to minimize the distance between points defined in each image and pose. In this case, we attempt to find a calibration that minimizes the distances between the circle, as described by equation 3.5, of each image. Using the full circle is obviously unnecessary, so one can define a subset of the circle based on a maximum distance away from the image plane. This maximum elevational distance is fairly important and will be discussed later. Since we now have a set of points representing the subset of the circle per image instead of a single point, we must now modify the standard algorithm used for point-based calibration.

$$\forall i = 1 \ldots n : \underset{X \in SE(3), c \in R^3}{\arg\min} (||c - B_i X p_i||) \tag{3.7}$$

Equation 3.7 is the standard approach for solving point-based calibration. B_i is the pose recorded by the external tracker, p_i is the point relative to the image, X is the unknown calibration transformation, and c is the unknown fixed point in the external trackers coordinate frame. Any BXp ultrasound calibration solver can be used to compute both c and X. The main change in our algorithm is that this becomes an iterative process as shown in the workflow diagram in figure 3.6.

The first step is to solve for c and X using traditional methods while assuming

that there is no out-of-plane uncertainty. We use a variant of the gradient descent solver described by Ackerman et al. [26] with a cost function, $|B_i X p_i - B_j X p_j| = 0$, that we minimize for every pair of indices. This provides an initial estimate of c and X that we use in the next step. For each image, we wish to select a new p_i that is not strictly assumed to have no elevation component. As was mentioned previously, there is a subset of points based on our model that includes the true p_i. The new p_i will be chosen such that the difference between c_i and $B_i X p_i$ using the current estimate of c and X is minimized. This procedure is similar to the find closest point procedure in algorithms such as iterative closest point (ICP) [51]. This new set of p is then used in conjunction with the original set of B to solve for a new c and X. These two steps repeat until X converges and its change in an iteration reaches some predefined tolerance level. The same algorithm applies to curvilinear transducer calibration. The only difference is that the predetermined subset of points is smaller due to the difference between equation 3.5 and equation 3.6. In theory, this means that the algorithm will converge much quicker when using a curvilinear transducer as opposed to a linear transducer.

3.5 Methods

As was mentioned previously, we explored two uses of an active point ultrasound calibration phantom. These respectively focused on solving US calibration while

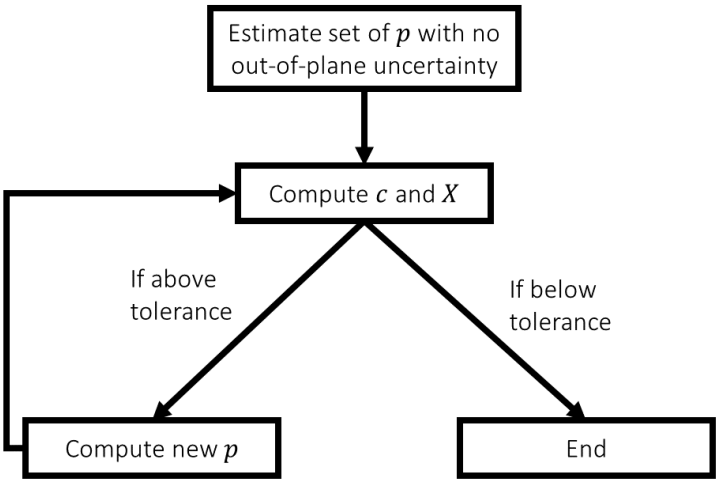


Figure 3.6: The workflow of the out-of-plane calibration algorithm is an iterative process that continually attempts to improve the estimation of c and X in one step, and the set of p_i in another.

localizing the PZT element within the ultrasound images mid-plane as well as solving ultrasound calibration while point targets are outside of the imaging plane. In this section, we describe the simulation and experiments conducted to evaluate these two uses.

3.5.1 Simulation Goals

We created a simulation using MATLAB based on the theoretical geometry shown in figure 3.5 under the assumption that the transmitters and receivers are ideal points. In our simulation, we observe the effects of changes on two parameters. The first is the amount of elevational distance that the point is from the US image plane. The y or out-of-plane component of each point is randomly chosen from a uniform distribution, bounded by the elevational distance parameter. The values of this parameter is

chosen to reflect the physical limitations of an ultrasound transducer. However, to simplify the simulation, we make this parameter independent from the ultrasound image depth, whereas it can be image depth-dependent in reality. The second is the standard deviation of the noise added to the distances between the active point and each of the transducer elements. If we relate this back to equations 3.4 and 3.5, we are adding noise to d_i. The added noise is chosen from a zero-mean Gaussian distribution with the standard deviation being the noise parameter.

Through these simulations, we want to show several things. First of all, we want to show that as the points used in US calibration move away from the US imaging mid-plane, this increase in error in the points elevational dimension will result in negative effects when using a conventional US calibration method. Showing this would indicate that a calibration phantom that allows for more accurate point localization is beneficial to US calibration. Second of all, we want to demonstrate that the out-of-plane US calibration algorithm is effective when points are not accurately localized within the US imaging mid-plane. To demonstrate these things and allow for a comparison with the conventional method, we will use the point reconstruction precision (RP) metric.

3.5.2 Experimental Setup

We placed an AE element and a CW phantom side by side at the same height in a water tank. The matching height allows for the data collection protocol described

in the next section. We acquire a 9cm depth US image from a 58.5mm L14-5W ultrasound probe (Ultrasonix Inc.) using the MUSiiC toolkit [52] to communicate with a Sonix Touch system. The UR5 robotic arm (Universal Robots Inc.), which has 6 degrees of freedom and an end-effector positioning repeatability of $\pm$100m, was chosen to be the external tracker and was rigidly holding the ultrasound transducer. The AE element is made of a customized PZT5H tube with an outer diameter of 2.08mm, an inner diameter of 1.47mm, and a length of 2mm. AUSPIS [50] is used to control the transmit and receive capabilities of the AE element. The CW phantom is made of two 0.2mm fishing lines. A fiber optical hydrophone (OH) developed by Precision Acoustics LTD was also added to the setup and placed at the same height to be used as a reference phantom. This device has a micro FabryProt acoustic sensor fabricated on a fiber tip [53]. It has a receiving aperture of 10m, a bandwidth of 0.25-50 MHz, a dynamic range of 0.01-15MPa, and a sensitivity of 150mV/MPa. The OH is used as a reference because it was previously shown to have better localization than the PZT AE points. Using the OH points as the test points allow us to further isolate the error in the calibration X from other errors.

3.5.3 Experimental and Data Analysis Procedures

The experimental data was collected using a strict protocol to allow for comparison between the CW, AE, and OH phantoms. The placement of the phantoms and how the data was collected can be seen in figure 3.7. The key point of this protocol was

attempt to standardize the tracking poses used for each phantom data set as much as possible. Since the phantoms were not located at the same position, complete standardization is impossible, so we fixed the tracking pose rotation. The robot motion was restricted to a single dimension such that each of the phantom point would be located at a similar region of the ultrasound beam transmission profile and the ultrasound image. This procedure is repeated for each pose until a total of 60 were acquired. We also repeated these experiments several times under different imaging conditions.

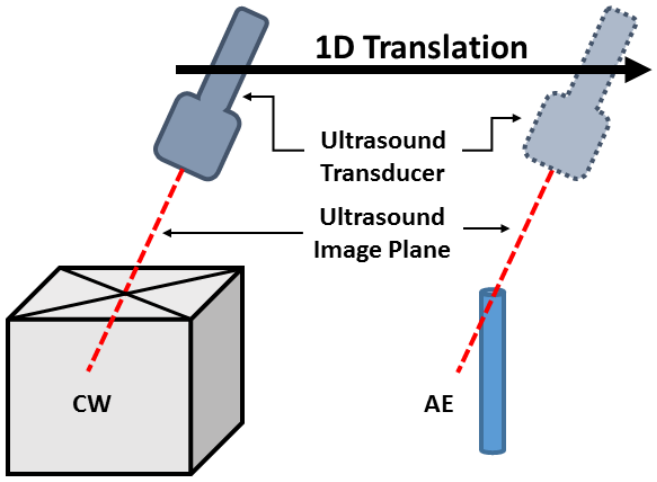


Figure 3.7: Experimental setup and data collection protocol to have comparable calibration data between CW and AE phantom. The ultrasound transducer is translated in one dimension to image the point from both phantom in a similar region of the image using the same pose orientation.

We analyze the experimental data using the RP metric. This metric is computed as shown in equation 3.8. X is computed with the calibration data set, while RP is computed using the test data set. This ensures that the test data is independent of the calibration data. Since we also manually segment these datasets, there is some

variance in the RP result for each phantoms depending on the user segmentation. We decided to show the best RP from these datasets. This is somewhat unfair to AE calibration, because it was shown by Guo et al. [22] that user segmentation of the CW point has more variance than user segmentation of the AE point.

$$\forall B, p \in Data_{test}, RP = norm(std(B_1 X p_1, \ldots, B_n X p_n)) \tag{3.8}$$

We also analyze the results of using automatically segmented AE points, as described in a previous section. This analysis was done by comparing the RP of CW, AE, and auto-AE calibration, while using the OH as the test data set. The aim is to show that automatically segmented AE points can have comparable results to expert segmentation. Some of this experimental data was also used to compute the RP when using the out-of-plane calibration method.

3.6 Results and Discussion

3.6.1 Ultrasound calibration simulation

Figure 3.8 shows the simulated reconstruction precision of a traditional BXp ultrasound calibration method when the calibration points are allowed to deviate from the imaging mid-plane by increasingly large magnitudes. As expected, we can see that the reconstruction precision increases as the elevational uncertainty increases. Thus,

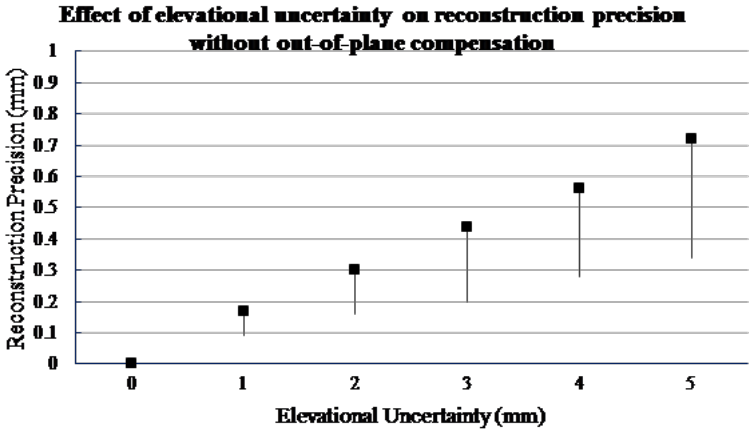


Figure 3.8: Effect of elevational uncertainty on reconstruction precision when elevational uncertainty in the calibration points increases without out-of-plane compensation

from a simulation standpoint, it is beneficial to use an active phantom for improved elevational localization.

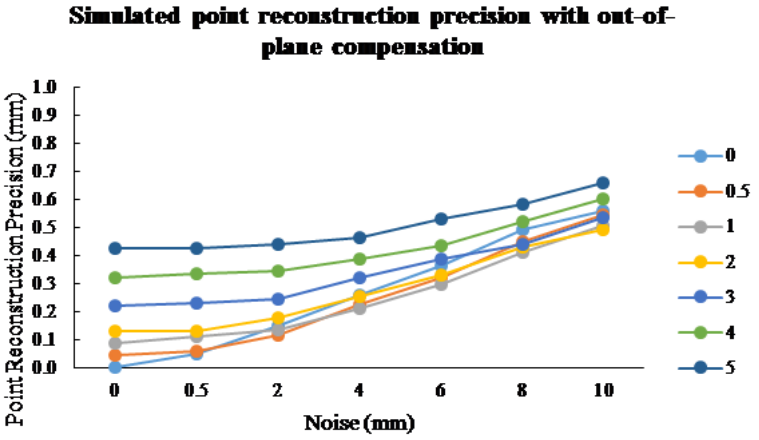


Figure 3.9: Simulated point reconstruction precision with out-of-plane compensation

Tables 3.2 and 3.1 and figures 3.9 and 3.10 show the point reconstruction precision for conventional and out-of-plane ultrasound calibration under different elevational

N=50	Noise=0mm	0.5mm	2mm	4mm	6mm	8mm	10mm
Ele=0mm	0±0	0.05±0.02	0.15±0.06	0.26±0.09	0.36±0.12	0.49±0.16	0.56±0.20
0.5mm	0.05±0.03	0.06±0.02	0.12±0.04	0.23±0.09	0.32±0.12	0.45±0.17	0.54±0.19
1mm	0.09±0.05	0.11±0.04	0.14±0.04	0.21±0.07	0.30±0.10	0.41±0.15	0.51±0.18
2mm	0.13±0.05	0.13±0.05	0.18±0.06	0.25±0.08	0.33±0.10	0.43±0.14	0.49±0.16
3mm	0.22±0.08	0.23±0.09	0.24±0.08	0.32±0.10	0.39±0.12	0.44±0.14	0.54±0.18
4mm	0.32±0.12	0.33±0.13	0.34±0.12	0.39±0.11	0.44±0.14	0.52±0.15	0.60±0.21
5mm	0.43±0.17	0.42±0.15	0.44±0.17	0.46±0.15	0.53±0.17	0.58±0.19	0.66±0.21

Table 3.1: Simulated point reconstruction precision with out-of-plane compensation under different elevational uncertainty and noise conditions (mm)

uncertainty and noise conditions.

There are several observations to make with respect to figures 3.9 and 3.10. First, the obvious observation is that if we look at these two figures independently, we can see that there is a general trend that both elevational uncertainty and noise will increase the point reconstruction precision. More interestingly, the detrimental effects of elevational uncertainty seem to decrease as noise increases. We can see this because the reconstruction precisions begin to bunch together when noise increases. This would seem to indicate that ultrasound calibrations are more sensitive to noise than elevational uncertainty. However, the noise range used in this simulation is also quite high. In practice, one would expect segmentation errors and noise to be less than a centimeter. The final observation to make is that the RP is generally lower when using the out-of-plane ultrasound calibration method. Similar to the previous

N=50	Noise=0mm	0.5mm	2mm	4mm	6mm	8mm	10mm
Ele=0mm	0±0	0.05±0.02	0.15±0.06	0.25±0.08	0.35±0.11	0.45±0.15	0.51±0.18
0.5mm	0.07±0.02	0.06±0.02	0.12±0.04	0.22±0.08	0.30±0.10	0.42±0.16	0.50±0.19
1mm	0.13±0.04	0.11±0.04	0.14±0.04	0.21±0.07	0.29±0.09	0.39±0.15	0.48±0.18
2mm	0.23±0.07	0.21±0.07	0.22±0.07	0.27±0.08	0.34±0.11	0.43±0.14	0.49±0.16
3mm	0.32±0.12	0.32±0.11	0.31±0.10	0.37±0.12	0.41±0.12	0.46±0.15	0.54±0.18
4mm	0.42±0.14	0.42±0.15	0.41±0.14	0.44±0.13	0.48±0.15	0.55±0.16	0.62±0.21
5mm	0.53±0.19	0.52±0.18	0.51±0.19	0.53±0.17	0.58±0.19	0.62±0.21	0.69±0.22

Table 3.2: Simulated point reconstruction precision without out-of-plane compensation under different elevational uncertainty and noise conditions (mm)

observation, the difference is more pronounced when the noise is lower.

3.6.2 Ultrasound calibration experiments

N=60	Point Reconstruction Precision(mm)							
Calibrate/Test Points	AE/CW	AE/Opt	AEOOP/Opt	CW/AE	CW/Opt	CWOOP/Opt	Auto-AE/Opt	Auto-AEOOP/Opt
Exp. 1	1.05			2.36				
Exp. 2	1.07	0.86		1.72	0.88			
Exp. 3	0.87	0.85		1.08	0.87		0.85	
Exp. 4		0.77	0.75		0.76	0.74	0.78	0.76

Table 3.3: Point reconstruction precision of X for best segmentation computed with calibration and test data.

A total of 4 experiments were conducted. While they were intended to compare specific methods, some broader conclusions can also be found across experiments. In the first experiment, the quality of the collected images is low for both the CW and

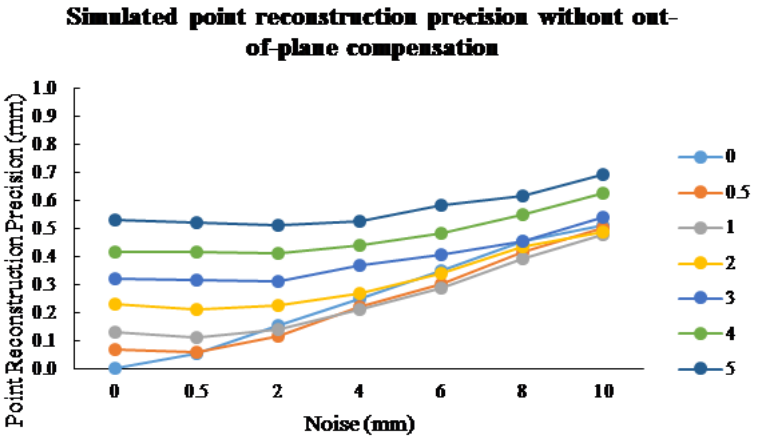


Figure 3.10: Simulated point reconstruction precision without out-of-plane compensation

AE phantoms, mainly due to multiple reflections and an unfocused beam. This data resulted in larger point reconstruction precisions as seen in table 3.3. In experiments 2, 3, and 4, the experimental setup is configured to improve the image quality. From the AE/CW and CW/AE columns in table 3.3, we can see that different image quality and subsequent segmentation under different conditions has a much larger effect on CW than AE phantoms.

Table 3.3, especially the columns using the OH as the test data set, shows that we can achieve comparable point reconstruction precisions using AE calibration and CW calibration. This contradicts our initial hypothesis of better point reconstruction precision. One possible reason for this may be the PZT element in our current AE setup. As we previously mentioned, the PZT element itself has a diameter of 2mm, which severely limits its mid-plane localization accuracy. This is due to there being uncertainty in both the elevational thickness of the transmission beam as well as the

origin of the maximum intensity received on the PZT element.

Table 3.3 also shows that the method with out-of-plane estimation has similar point reconstruction precision to the method without out-of-plane estimation. This could be an indication that the point sets used in these experiments were already fairly well-positioned within the ultrasound image mid-plane. This is also similar to the situation shown in simulation, where the method has little effect when the elevational uncertainty is low. Based on these observations, this method may be more beneficial when applied towards data sets collected by those with less calibration experience.

3.6.3 General discussion and future improvements

From our experimental result, we also see that we had a larger error than in simulation. While this is mostly expected, one very likely reason is that the assumptions that we originally made in our algorithm and simulation are not quite correct. The ultrasound transducer and active point are not ideal point sources in reality. Since the equations previously assume otherwise, we naturally expect there to be deviations between simulation and reality.

With that being said, there are several possible improvements to the active phantom. First, the active element would ideally be as small as possible. It is believed that this would improve the mid-plane localization accuracy and the results using the OH also reinforce this belief. In addition, one drawback of the current active phantom is that there is no imaging or electrical feedback unless the US imaging plane is very

close to the phantom. An obvious extension is to combine the active phantom with a CW phantom, using the wires to guide the user when the US transducer is far away and using the active phantom for fine adjustments.

An observation that we had over the course of our experiments was that this out-of-plane calibration algorithm can be prone to over-fitting. As we previously mentioned, we select a subset of the circle with some maximum elevational distance. Increasing this distance will almost always allow for a better least squares fit. Thus, one needs to apply caution when picking this parameter and should choose it to fit the actual experimental scenario. We currently apply a static parameter to all of the images in our dataset, but one can envision an algorithm which uses a variable parameter based on some other notion of distance away from the ultrasound image mid-plane. One example of feedback that can facilitate this would be signal intensity, as it decreases the further away the active point is from the mid-plane. Another possibility is to use the active echo feedback described by Guo et al. [50].

The two calibration methods may seem contradictory at first, but they can complement each other greatly. As we mentioned previously, AE ultrasound calibration allows the point target to be more accurately localized within the imaging plane, but there will still remain some error. Out-of-plane ultrasound calibration can account for this uncertainty. At the same time, fine localization can be extremely time-consuming. With the out-of-plane approach, we can now approximately place the active point in the ultrasound mid-plane and estimate for out-of-plane deviations. The feedback is

still useful, but fine-tuning the position generally takes proportionally more time than getting to the general area. While this out-of-plane approach can theoretically account for points of any out-of-plane distance, the preferred embodiment is to attempt to place the active point in the ultrasound mid-plane and to use this out-of-plane approach as a slight adjustment to avoid over-fitting.

3.7 Conclusion

In this chapter, we discussed and demosntrated multiple uses of an active phantom for the purposes of ultrasound calibration. We were able to show that active phantoms can lead to more image configuration independent ultrasound calibrations. We also showed that the fully automatic segmentation method can achieve the same point reconstruction precision as manual segmentation. Finally, we presented an out-of-plane ultrasound calibration method and showed its feasibility through simulation and experiment. Future work will include an optimized AE element design for calibration applications.

Part II

Tracking Ultrasound Transducers in an Interventional Photoacoustic Surgical System (I-PASS)

This part presents the work done towards realizing the system shown in figure 4.1B. The main objective is to establish direct registration between US imaging and another imaging modality such as video or computed tomography. Several surgeries require real-time US including liver resections, partial nephrectomies and prostatectomies and real-time fusion of US and video is crucial to their success. Our approach is to create virtual photoacoustic fiducials, made of light, at the air-tissue interface. A projection system will be used to project these photoacoustic fiducials onto the surface of the organ through air. At the air-tissue interface, these landmarks can be seen both in US with the photoacoustic effect and in video. The following chapters explore the use of these photoacoustic fiducials under different scenarios with different ultrasound transducers. Chapters 4 and 4 look at the use of three-dimensional ultrasound transducers. Chapter 6 looks at the use of two-dimensional curvilinear ultrasound transducers. Chapter 7 looks at the use of bi-planar trans-rectal ultrasound transducers.

Chapter 4

Three-dimensional ultrasound to video registration with sequential photoacoustic markers

4.1 Introduction

Interventional guidance systems are becoming increasingly common in modern surgical procedures including open, laparoscopic, and robotic surgeries [54]. During such procedures, surgeons can lose track of tumors as they move in and out of the cameras field of view. Guidance systems can be used to alleviate these concerns by providing a fusion of video with other imaging modalities, such as intraoperative ultrasound (US), to aid the surgeon in locating tumors or other regions of interest.

The clinical utility of these guidance systems depend on the registration of other surgical tools and devices with the guidance system, such as stereoscopic endoscopes and US transducers.

The registration between US images and video visualization remains a significant challenge. Typically, electromagnetic (EM) or optical navigational trackers [18, 55] are used to provide real-time position and orientation about tools such as US probes and endoscopic cameras. These navigational trackers usually track sensors or markers relative to a separate base station placed within the surgical setting, adding complexity. However, this approach has serious limitations and is subject to error buildup from multiple tracking and calibration errors.

EM-based surgical navigation and tracking systems [18, 56] are the most common choice for laparoscopic surgery, flexible endoscopy, and other minimally invasive procedures because a clear line of sight is not required between the base station and the attached sensors (see figure 4.1A). However, there are several drawbacks associated with using an EM-based system. First, wired EM sensors must be placed on the US transducer before it can be tracked by the surgical navigation system. This could decrease the surgeons comfort while potentially increasing the cost associated with handling and sterilizing modified surgical tools. Second, a large intrusive EM base station must be placed in close proximity with the operating table, adding clutter to the surgical setting. Third, EM-based systems also suffer from magnetic field distortions when metallic objects are placed within its field. This drawback is particularly

significant as the tracked sensors will have significant errors when this is the case.

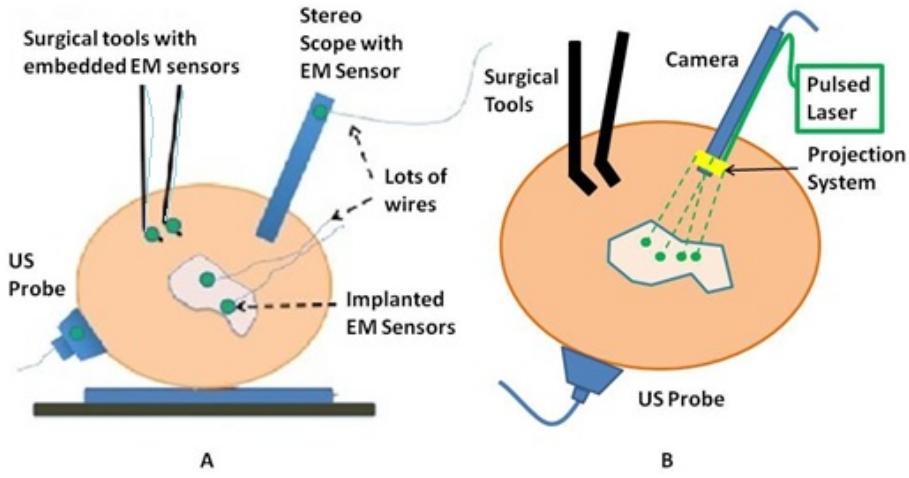


Figure 4.1: A) Standard EM-based Navigation System, B) PA Navigation System [R21 NIH-NIBIB-EB015638]

Optical tracking systems such as those developed by Claron Technology Inc. (Toronto, Ontario, Canada) or Northern Digital Inc. (Waterloo, Ontario, Canada) avoid the field distortion problems associated with EM trackers and frequently do not require wired sensors. Studies have shown that optical trackers can detect their optical markers with sub-millimeter accuracy [57, 58], but line-of-sight restrictions often make them impractical for laparoscopic procedures. Markers can certainly be placed outside the body, but this will degrade the tracking accuracy for long and flexible tools inserted into the body.

Another drawback for both optical and EM-based systems is the need to acquire the transformation from the tool to the surgical navigation system indirectly, i.e. the transformation of interest is computed via a chain of transformations over several coordinate systems. An example of this can be seen in 4.2. This case applies to both

EM and optical tracker-based systems. As an example, a transformation of interest would be one that goes from the EM base-stations coordinate system to the US images coordinate system. To acquire this transformation, one must obtain both the transformation from the EM base-station to the EM marker and the transformation from the EM marker to the US image. The first transformation is given by the tracking information from the EM base-station and the second must be obtained through a calibration process. Calibration is a topic where many authors have presented research to achieve better accuracy and lower errors [41, 59, 60]. Their results have shown that the calibration process dominates the overall error in the registration [41, 57, 58]. Other studies have presented overall registration errors of approximately 1.7-3mm for artificial phantoms and approximately 3-5mm for tissue [18, 56, 61, 62].

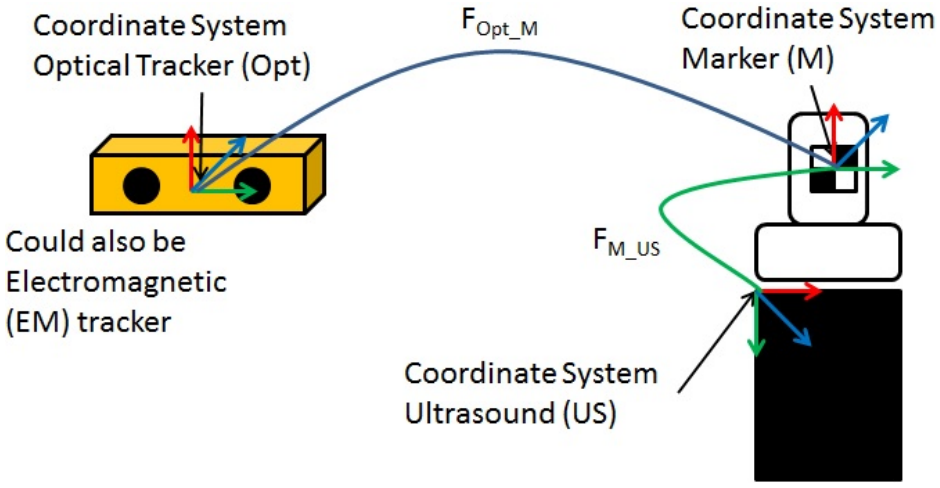


Figure 4.2: An example of the indirect chain of transformations necessary to enable interventional US guidance.

Yip et al. [62] demonstrated a registration method that uses a tool at the air and tissue boundary. This tool had optical markers in the stereo camera (SC) space on

one side and US-compatible fiducials on the other. A drawback of this method is that it requires a custom registration tool to be in direct contact with the tissue. Also, the US fiducials must be segmented from US B-mode images. This can be a difficult task as the speckle or structural information contained in US B-mode images may obscure the US fiducials.

Vyas et al. [63] demonstrated proof of concept for a direct registration method with the photoacoustic (PA) effect requiring a single transformation between the frames of interest as opposed to a chain of transformations. This method addresses each of the drawbacks above. Markers or sensors are not necessary to generate a coordinate transformation between the tracker frame and the US transducer frame so the tools that surgeons use will remain the same. Previous work [10, 64] showed that a pulsed laser source can effectively generate a PA signal in tissue, resulting in an acoustic wave that can be detected by conventional US transducers [65, 66]. Other than the US transducer, nothing else needs to touch the surface. Each laser point projection was seen as a green spot in the SC space and as a PA signal in the US space. Segmentation of the PA signal is also simpler in a PA image than a US B-mode image because the laser spot was now the only acoustic source. Finally, the calibration process is unnecessary since the coordinate transformation from the SC frame to the US frame can be computed directly with the two 3D point sets based on rigid registration algorithms [44, 55].

This chapter examines a direct 3D US to video registration method and demon-

strates its feasibility on *ex vivo* tissue. Improving on the work of Vyas et al. [63], we used a 3D US transducer instead of a 2D US transducer to detect the PA signal. Using a 3D transducer allows this registration method to function for a non-planar set of 3D points. This is a significant requirement because we aim to deploy this method in a laparoscopic environment and organ surfaces will rarely form a planar surface. We also improved significantly on the point-finding algorithms used by Vyas et al. [63] to find the PA signal in both SC images and US volumes. In addition to using a synthetic phantom with excellent light absorption characteristics, we also used resected *ex vivo* porcine liver, kidney, and fat each individually embedded in gelatin phantoms to demonstrate this methods feasibility for the eventual guidance of laparoscopic tumor resections and partial nephrectomies. These phantoms are representative of our proposed clinical scenario since the laser light will likewise only interact with the surface of the phantoms. The gelatin acts purely as a support material and does not affect the PA signal generation. This section provides more detailed information about our processing methodology, a demonstration of the ability to generate a PA point signal, an analysis of the point localization errors in various phantoms, and registration error results on multiple *ex vivo* tissue phantoms.

This chapter details the technical approach, experimental methods, and three key results: the ability to generate a PA point signal, a comparison of the point localization errors in the SC and US domains, and registration error results. We will discuss the significance of our results, potential errors, and a detailed roadmap to

eventual implementation in a surgical setting along with future directions.

4.2 Contributions

The main contributions of this chapter are the following:

- Development of a photoacoustic-based tracking method for three-dimensional ultrasound transducer
- Experimentation on phantoms and *ex vivo* tissue to validate this tracking method
- Demonstration of sub-millimeter target registration errors when using this tracking method

4.3 Acknowledgement

I wish to thank Xiaoyu Guo for help with the optical setup, Hyun-Jae Kang and Nathanael Kuo for help with the MUSiiC toolkit, and Daniel Carnegie for help with *ex vivo* phantoms. I was responsible for leading all components of this work.

4.4 Technical Approach

This approach can be split into a data collection phase, a data processing phase, and a registration phase. The data collection phase outputted SC image pairs, five images for each camera, and a 3D RF US volume for each projected laser spot.

The data processing phase used the data and generates a 3D SC point set and 3D US point set. Finally, the registration phase used the two point sets to compute a coordinate transformation from the SC frame to the US frame. Figure 4.3A shows the experimental setup using the porcine liver phantom and an overlay of a US image representation using the inverse of the computed transformation.

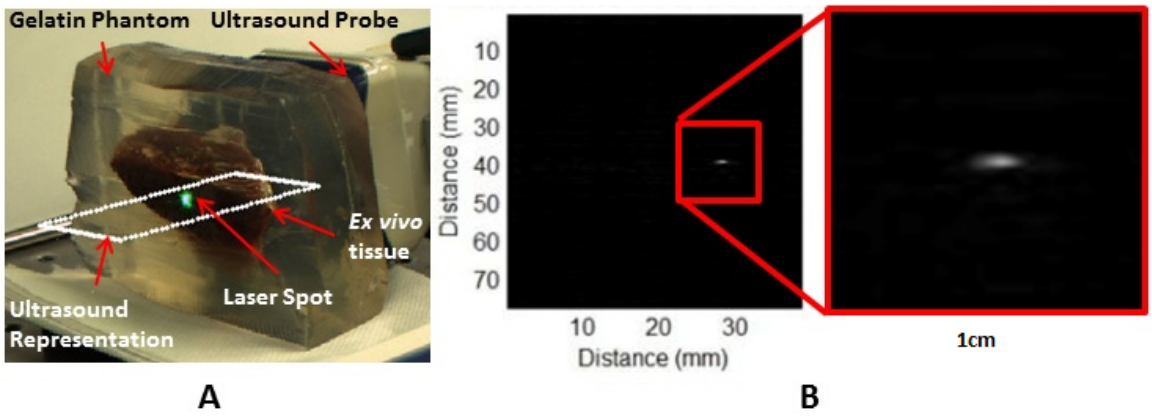


Figure 4.3: A) Experimental Setup and Video Overlay, B) PA Signal within an US image

4.4.1 Data Collection

Figure 4.4A shows the workflow of the data collection phase of our experiments. First a laser spot is projected onto the exposed surface of the *ex vivo* tissue, *ex vivo* fat, or synthetic material. It is important to emphasize that most of the laser energy from these laser spots is absorbed at the surface of the phantom. There are inaccuracies in SC spot triangulation if the laser spot is projected at or near the tissue or fat gelatin interface because the laser spots become irregularly shaped when projected

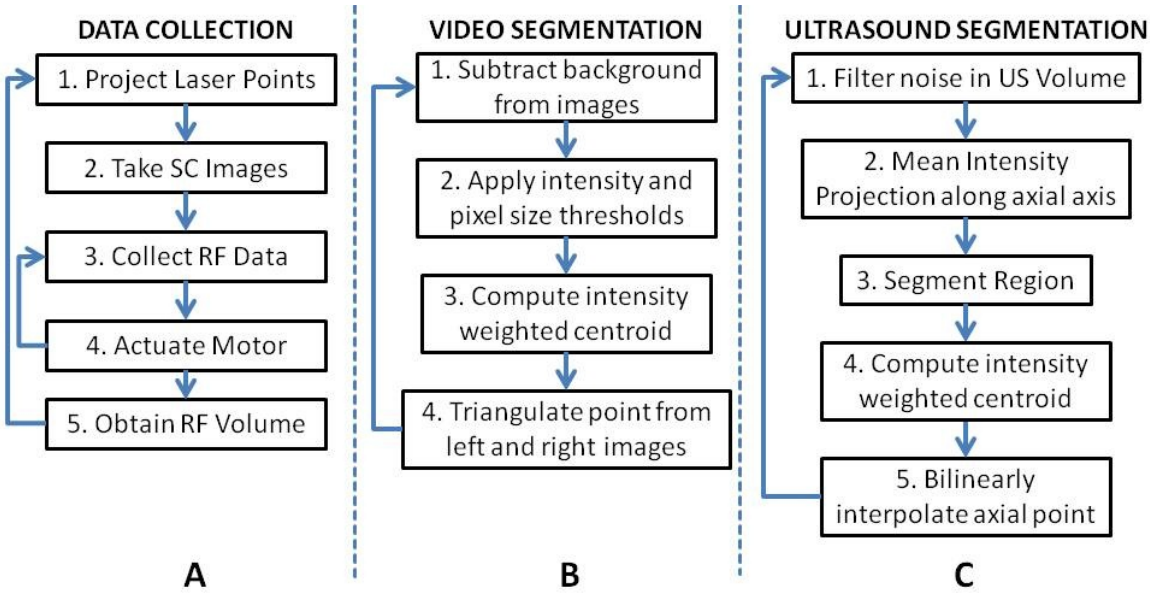


Figure 4.4: Workflow for A) Data Collection, B) SC Segmentation, and C) US Segmentation

onto clear materials. When projecting onto the fat, there was significant reflectance and saturation due to its color. To reduce these effects, we placed laser goggles in front of the cameras, acting as a filter that modulates light intensities with varying wavelength dependent optical densities. Second, several images were taken with each camera. The laser spot projected onto the phantom must be visible in at least one image per camera for triangulation to be possible. Our cameras had a faster capture rate than our lasers repetition rate, so some of the frames were devoid of the laser signal. We exploit this during data processing. Steps 3 and 4 show that the 3D US transducer motor actuation and RF data were intermittently collected from the DAQ device to scan and acquire the RF data of the volume of interest. The motor step size was 0.49°. The volumes field of view varied amongst the experiments because the

phantoms were different sizes and several of them did not require as many slices to generate a volume that covered the entire phantom. The probe has a lateral length of 38.4mm and the resulting PA images have a depth of 8cm. A real-time implementation is feasible and an automatic process is in development. Currently the manual process takes approximately 3 seconds per slice. This workflow was repeated for each of the laser spots.

From the data collection phase, we present results showing that we can generate a PA point signal on the various phantoms using 532 nm, 1064 nm, or both. In each of these cases, the laser energy is mostly absorbed at the surface and the laser light does not interact with any medium other than air and the surface of the phantom. As we described, only a small amount of energy per pulse is required to generate a single PA spot and to localize it at the air-tissue interface. These are shown as a series of images similar to the one shown in figure 4.3B. The size or location of the PA spot is unrelated between scenarios as parameters such as spot size or laser energy were different. Since we are displaying a 2D slice from a 3D volume, it is also possible that the displayed image does not contain the centroid of the PA point signal.

4.4.2 Data Processing

The data processing phase involved the segmentation of the SC images into 3D SC points, the segmentation of the 3D RF US volume data into 3D US points, and the computation of the transformation from the SC frame to the US frame.

Figure 4.4B shows the workflow for SC segmentation. For each camera, we picked a SC image with the laser spot and without the laser spot. Next, the background images without the laser spot were subtracted from the images with the laser spot. As we can see from figure 4.5A, the laser spot is nearly segmented from the scene. We used an iris to decrease the beam size of the laser and the reflection from the iris results in the laser spot in the lower left. To compensate for this spot, we neglected an appropriate border on each image. We then applied an intensity and pixel size thresholds such that the laser spot is segmented out. These thresholds were selected based on the laser beam diameter and the phantoms reflectance and are varied between the different scenarios. The thresholds were selected manually, but a method to automate threshold selection based on experimental parameters is in development. Next, we fitted an ellipse to the segmented region and computed the intensity weighted centroid based on the pixels within the ellipse from the original image, resulting in the image shown in figure 4.5B. Calibration files for our specific SC allowed us to triangulate the segmented point from each camera and to obtain a single 3D point in the SC frame. In our current implementation, this step takes approximately 500ms for each laser spot projection. This workflow was repeated for each laser spot projection.

The workflow for the segmentation of the 3D RF US volume is shown in figure 4.4C. First, for each slice of a 3D RF US volume, the RF data was beamformed using the k-wave toolbox [67] in MATLAB. The dynamic range of the image was normalized with respect to the volume and we applied a threshold to decrease the size of the PA

Figure 4.5: Resulting Images of Video Segmentation Workflow for A) Step 1, B) Step 3

signal seen in each volume. These thresholds ranged from 0.4 to 0.5. The size of the PA signal refers to the number of non-zero valued pixels representing the PA signal after thresholding. A smaller size leads to a more compact representation of the PA signal, but it is also important to maintain the characteristic elliptical shape of the PA signal. Figure 4.3B shows a k-wave beamformed PA signal image. Beamforming requires approximately 140ms for each PA image. Next, we projected the volume onto the lateral-elevational plane by taking the mean along each axial ray. An intensity and pixel size threshold were then applied to this image. These thresholds were selected in a similar fashion to the ones used for SC segmentation. An ellipse was fitted on the segmented region and an intensity-weighted centroid was computed resulting in lateral and elevational coordinates. Figure 4.6A is an example of this step showing the lateral-elevational image and the corresponding ellipse. As described earlier, the PA signal originated from the surface and any

penetration into the surface. Since air cannot generate a PA signal in our setup, we exploit that the high intensity pixels farthest away in the axial direction are from the surface. Thus, we obtained the axial coordinate corresponding with a lateral-elevational coordinate as the axial-most high intensity pixel. This step is particularly important because the penetration of the laser pulse was deeper for *ex vivo* tissue than the penetration for the synthetic phantom because the laser energy was not completely absorbed at the surface. We used bilinear interpolation to obtain axial coordinates between sampled points. These three coordinates were converted to 3D US coordinates based on transducer specifications. The lateral coordinate combined with the lateral resolution of our US transducer results in the lateral coordinate in 3D US space. The axial coordinate combined with the axial resolution represents a ray protruding from the US transducer, while the elevational coordinate relates to the angle this ray makes with the first US image in the volume. Solving this geometry problem results in a 3D US coordinate. The computation time for this process, not including beamforming, is correlated to the number of slices and field of view in each volume. Thus, the synthetic phantom required approximately 730ms per volume and the other phantoms required approximately 570ms. A 3D RF US volume was acquired for each PA spot.

From the data processing phase, we present the localization accuracy of the US PA signal segmentation. In these results, we assume that the segmented SC points are the ground truth. For each US point set, we compute the distance between each

pair of points. We compare the US distance of a particular pair of points with the distance between the same pair of SC points. The average of all differences for all point pairs is the recorded localization error. These are reported in each of the lateral, axial, and elevational axes and the overall Euclidean distance for each experiment.

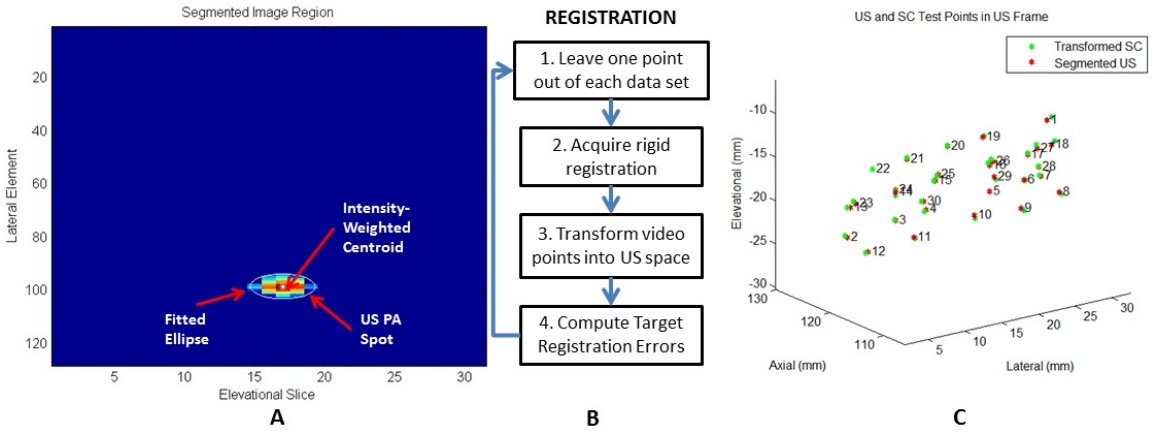


Figure 4.6: A) Resulting Image of US Segmentation Workflow Steps 2, 3, and 4 B) Workflow for Transformation C) US and SC Point Cloud Registered Together

4.4.3 Registration

The transformation from the SC frame to the US frame was computed with the 3D SC and 3D US point sets as shown by the registration workflow in figure 4.6B. Any registration method for computing the transformation between two 3D point sets can be used. We used the coherent point drift algorithm [21] in our experiments. One of the main reasons for using coherent point drift was that it allows for data points to be missing from either dataset. An assumption that we have made is that every laser spot will be visible in the SC images and every PA signal will be visible in the

US volume. This assumption was valid for our experiments, but may not hold in the surgical setting due to SC or US transducer movement. Some of the points may lie outside of the field of view of either the SC or US transducer. The coherent point drift registration algorithm allowed us to acquire a registration as long as there were enough corresponding points in the SC images and the US volume. More experiments will be necessary to determine how many points are enough and if there will be any corresponding tradeoff in accuracy.

In figure 4.6C, each SC point was independently used as a test point while the rest of the points in the point set were used to transform the SC test point into the US domain. The figure shows the test SC point alongside its corresponding US point with a vector connecting the two. Our results from the registration workflow are shown as a series of target registration errors (TRE) computations in each of the lateral, axial, and elevational axes and the overall Euclidean distance for each experiment.

The transformation from the SC frame to the US frame was used to transform the 3D SC points to the US frame for validation. The inverse transformation was used to display a representation of an US image into the SC frame as shown in Figure 4.3A.

4.5 Methods

In these experiments, we used a Q-switched neodymium-doped yttrium aluminum garnet (Nd:YAG) Brilliant (Quantel Laser, France) laser to generate a PA marker on

various materials. We used different combinations of wavelengths (532nm, 1064nm) and energy densities (˜6mJ/cm^2, 19mJ/cm^2, 64mJ/cm^2, 172mJ/cm^2, and 45mJ/cm^2) with specifics for each experiment indicated in the results section. These values do not represent the lowest possible energy necessary to generate a PA signal on our materials. We chose these values to give our PA images a sufficient signal-to-noise (SNR) ratio without having to average over multiple frames. The SNR for some sample images can be seen in table 4.1. The images are first normalized then the SNR is computed as the mean of the foreground divided by the standard deviation of the background, where the foreground and background is separated by a threshold described later. It should be noted that the maximum permissible exposure (MPE) is 19.5mJ/cm^2 for 532nm and 97.5mJ/cm^2 for 1064nm as calculated from the IEC 60825-1 laser safety standard [68] based on a 0.25s exposure time, a 4ns pulse width, and a frequency of 10Hz. We used a SonixCEP US system and a 4DL14-5/38 US transducer developed by Ultrasonix Medical Corporation (Richmond, Canada) to scan the volume of interest. This US transducer has a motor that actuates a linear US array to move angularly around an internal pivot point. This US transducer has a bandwidth of 5-14 MHz and the linear array is approximately 38mm in length. The Sonix DAQ device, developed by the University of Hong Kong and Ultrasonix, and the MUSiiC toolkit [52] are used to acquire pre-beamformed radiofrequency (RF) data from the US machine. The k-wave toolbox [67] in MATLAB (Mathworks Inc. Natick, MA) is used to beamform and reconstruct PA images based on the pre-beamformed

RF data. For our SC setup, we used a custom system containing two CMLN-13S2C cameras (Point Grey Research, Richmond, Canada) to capture images at 18Hz. A camera calibration process using the Camera Calibration Toolbox for Matlab [[69] generates a calibration file for our SC setup. These calibration files contain the SC setup intrinsic parameters to do 3D triangulation. We created several phantoms for these experiments: A synthetic phantom made with plastisol and black dye, an *ex vivo* liver phantom made with gelatin and a freshly resected porcine liver, an *ex vivo* kidney phantom made with gelatin and a freshly resected porcine kidney, and an *ex vivo* fat phantom made with gelatin and porcine fatback. The surface of the *ex vivo* tissue or fat is exposed and not covered by the gelatin.

4.6 Results

Three sets of results from various points in our experiment are presented: the ability to generate a PA point signal, a comparison of the point localization errors in the SC and US domains, and registration error results.

The first set of results indicates that we can generate a PA point signal on a variety of materials using various energy levels or wavelengths. Table 4.1 specifies each of the five scenarios that we tried. Figure 4.7 shows the US images displaying a PA signal corresponding with each of the scenarios in table 4.1.

The second set of results includes the localization errors of the PA points in

Scenario	1	2	3	4	5
Phantom	Synthetic	Liver	Kidney	Fat	Fat
Wavelength	532nm	532nm	532nm	532nm	1064nm
Energy Density	6mJ/cm^2	19mJ/cm^2	64mJ/cm^2	172mJ/cm^2	45mJ/cm^2
MPE	19.5mJ/cm^2	19.5mJ/cm^2	19.5mJ/cm^2	19.5mJ/cm^2	19.5mJ/cm^2
SNR	3.5	2.7	4.1	3.6	3.8

Table 4.1: Observed Laser Energy Densities in Different Scenarios

scenarios one to four outlined in table 4.2. Scenario five is not included because 1064nm wavelength light is invisible to the SC. This means that scenario five is not applicable to our experiments with the current design. The metric that we use is defined in equation 4.1. As mentioned previously, we compute the difference of the distance between a point pair in the US space versus the distance between a point pair in the SC space. This metric treats the SC points as the ground truth. The reported localization error (LE) is the average of these differences over all point pairs for each experiment.

$$LE = \frac{2!(N-2)!}{N!} \sum_{i=1}^{N} \sum_{j=1}^{N} (||SC_i - SC_j|| - ||US_i - US_j||) \tag{4.1}$$

The registration results of our experiments on the synthetic phantom, the *ex vivo* liver phantom, the *ex vivo* kidney phantom, and the *ex vivo* fat phantom are validated using the TRE metric defined in equation 4.2. F_{SCUS} is the transformation from the

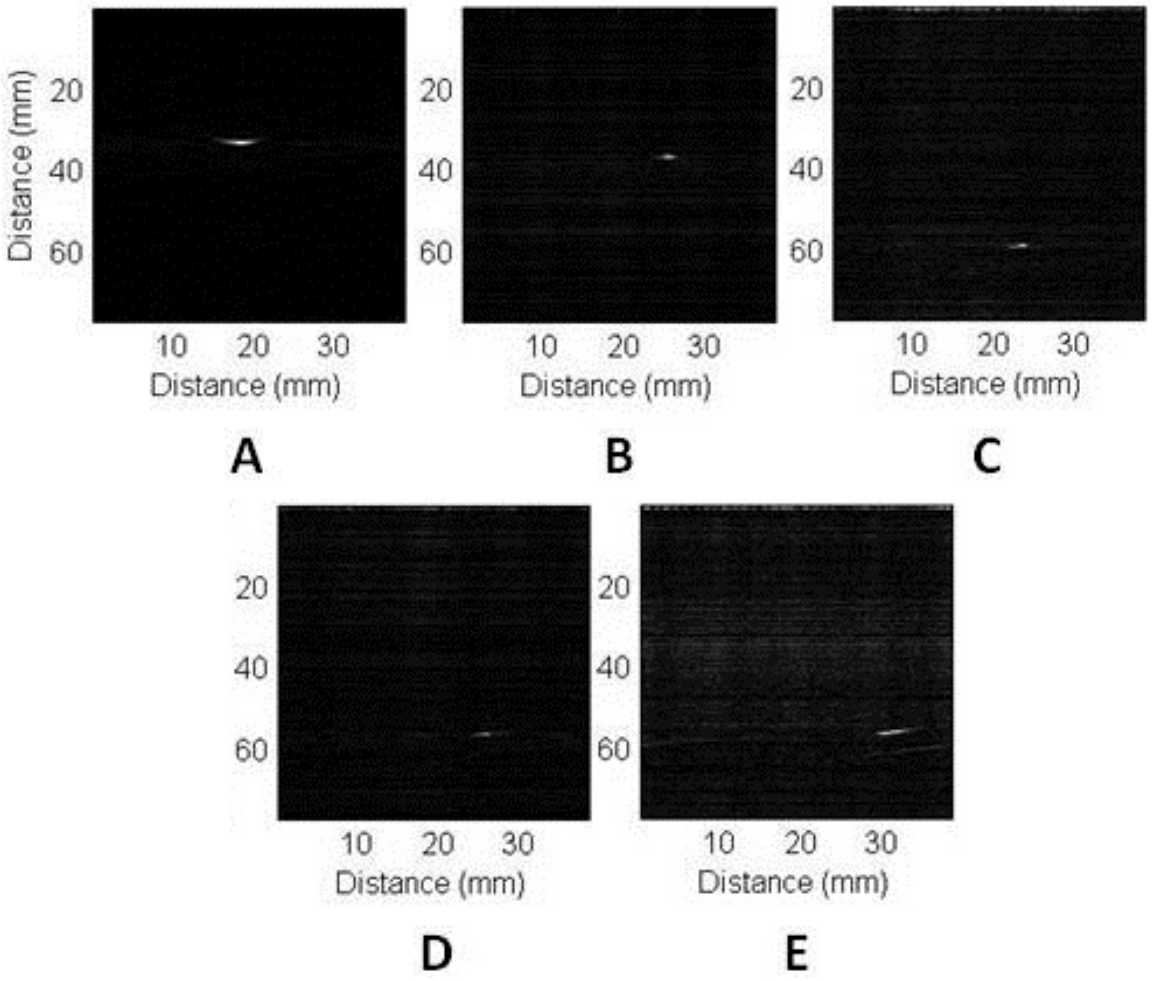


Figure 4.7: Sample PA Images for A) Scenario 1, B) Scenario 2, C) Scenario 3, D) Scenario 4, E) Scenario 5

SC frame to the US frame and is computed with all of the SC and US points except for one. The TRE is the difference between the actual US test point and the transformed SC test point in the US frame.

$$TRE = \frac{1}{N}\sum_{k=1}^{N}(||F_{SCUSk}SC_k - US_k||) \tag{4.2}$$

N is the number of points in a specific experiment and N-1 points is used to

Phantom	Synthetic	*Ex vivo* Liver	*Ex vivo* Kidney	*Ex vivo* Fat
Number of Point Pairs	435	435	190	55
Localization Error (mm)	0.46 ± 0.33	0.32 ± 0.23	0.19 ± 0.15	0.45 ± 0.39

Table 4.2: Average LE Results for Experiments

compute the transformation from the SC frame to the US frame. The remaining point is used as a test point to compute the TRE. This computation is repeated with each of the N points as test points. Table 4.3 shows the average and standard deviation of the TRE results for the N cases in the synthetic phantom, the *ex vivo* liver phantom, the *ex vivo* kidney phantom, and the *ex vivo* fat phantom experiments respectively. Figure 4.8 shows the TRE results in a box-whisker plot.

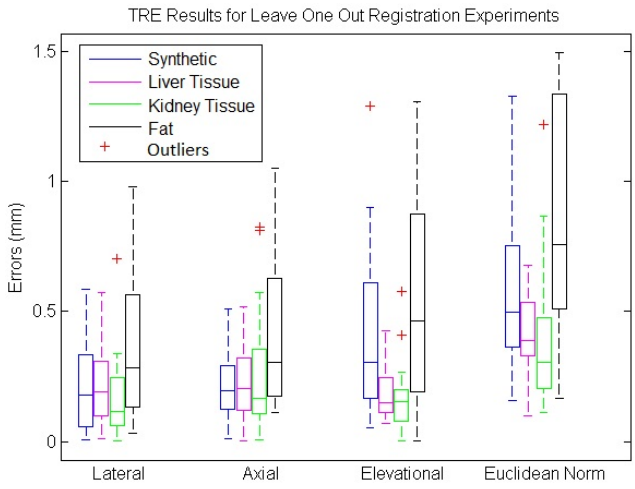


Figure 4.8: Box-Whisker plot of TRE Results for Leave One Out Registration Experiments

Phantom	Synthetic	*Ex vivo* Liver	*Ex vivo* Kidney	*Ex vivo* Fat
Number of Points (N)	30	30	20	11
Field of View °	19.6	14.7	14.7	14.7
Lateral (mm)	0.21 ± 0.17	0.22 ± 0.16	0.17 ± 0.16	0.36 ± 0.28
Axial (mm)	0.21 ± 0.13	0.24 ± 0.15	0.26 ± 0.24	0.43 ± 0.30
Elevational (mm)	0.41 ± 0.31	0.18 ± 0.10	0.16 ± 0.14	0.54 ± 0.42
Euclidean Norm (mm)	0.56 ± 0.28	0.42 ± 0.15	0.38 ± 0.28	0.85 ± 0.45
Min Norm (mm)	0.15	0.10	0.11	0.17
Median Norm (mm)	0.56	0.39	0.33	0.76
Max Norm (mm)	0.94	0.68	1.23	1.50
Registration Runtime (ms)	163	150	190	173

Table 4.3: Average TRE Results for Leave One Out Registration Experiments

4.7 Discussion

As seen in table 4.1, the energy densities used in the *ex vivo* liver and kidney experiments are close to or exceeding the MPE at that laser wavelength. These energy densities do not present a concern as we were not using the lowest possible energy density to generate the PA effect. Thus energy density levels below the MPE threshold are quite feasible. Additionally, averaging multiple US PA images at a lower laser energy density can allow us to retain the same SNR at the expense of time. The energy density at 532nm used in the *ex vivo* fat experiment does present

a legitimate concern as it was an order of magnitude higher than the MPE. Fat is also the most likely material encountered in a realistic surgical setting as it covers the organ of interest and surgeons try to remove as little of it as possible. One situation where this method may benefit from an *in vivo* setup as opposed to an *ex vivo* setup is the presence of blood. At a wavelength of 532nm, blood has a significantly higher absorption coefficient than fat [70, 71]. It is therefore a possibility that the blood on the fat surface be used to generate the PA signals as opposed to the fat itself. This would also mean that a much lower laser energy density is necessary.

As seen in table 4.1, the energy density at 1064nm used in the *ex vivo* fat experiment was beneath the MPE threshold. While we avoid concerns regarding energy density levels, this presents a situation where the SC used must be receptive to 1064nm light. This is usually not the case as cameras typically have a visible range of 400nm to 900nm. A possible solution is to project a continuous wave, low power, visible laser that is coincident to the 1064nm laser. However, this solution presents a new source of error as it may be difficult to achieve coincidental points. Another possible solution is to use a wavelength that is within the visible range of typical cameras, yet has a higher absorption coefficient for fat than that observed at a wavelength of 532nm.

The results in table 4.2 imply that the US PA spots are being localized fairly accurately. The average and standard deviation of the difference in distance between all point pairs in the US and SC domain are sub-millimeter for each experimental

scenario. These errors are in line with the point source localization errors for typical SC systems [57,58]. Future studies for determining the localization error may require a more accurate representation of ground truth data. A possible study is to project the laser spots on specific known locations of the phantom.

At the level of error measurements shown in table 4.3, it is likely that the calibration of the SC system is a significant contributor. They are able to track optical markers at sub-millimeter accuracy [57,58], so this error is usually negligible in comparison with the approximate 3mm errors from calibration [41, 57, 58]. Since our results were 0.56mm, 0.42mm, 0.38mm, and 0.85mm errors respectively, the SC systems error became significant. We used a custom SC system, so its errors were also likely greater than a commercial SC system.

From the experimental results shown in table 4.3 and figure 4.8, it can be seen that our system achieved sub-millimeter TRE measurements for each of our experimental scenarios with different phantoms. We wish to highlight that these results are significantly better than the overall registration errors of approximately 1.7-3mm for artificial phantoms and approximately 3-5mm for tissue [18, 56, 61, 62] presented in literature. There are several differences in the results between each scenario. First, the synthetic phantom had a larger Euclidean error than the *ex vivo* liver and *ex vivo* kidney phantoms almost entirely due to the elevational error. This was likely due to the larger field of view in the synthetic phantom experiment as well as normal variation across experiments. More experiments must be performed to obtain an average

error across multiple experiments. The *ex vivo* fat experiment had noticeably worse results in both the mean and standard deviation in each direction and the Euclidean norm. A possible reason for this is the light delivery system. To achieve the energy density necessary to generate a PA signal on fat, we needed to focus our laser beam. Our setup was such that the focusing of the laser beam was not uniform across all points. As mentioned before, laser goggles were also placed in front of the SC for this particular experiment. These two circumstances differ from the other scenarios and may have introduced inconsistencies in the SC or US point sets. Another possibility may be the smaller sample size of this experiment compared to the rest.

There are several considerations when discussing this systems deployment in our intended applications of laparoscopic tumor resections. The first is the placement of the transducer. In our experiments, we used a relatively large 3D US transducer that would be difficult to place inside the body during a laparoscopic procedure. However, the transducer is often placed externally [18,62] in these procedures, so the size of the probe is not an issue. Naturally, there are disadvantages of placing the transducer externally and farther from the region or organ of interest. The SNR of ultrasound images degrades as the depth increases, which would likely lead to errors in localizing fiducials or, in our case, the PA signal. However, since the PA signal only has to travel in one direction, as opposed to traditional US, our PA images will have better quality than US images of equivalent depth.

Another issue with our 3D US transducer was the acquisition speed. There are

certain applications where an acquisition speed of a volume per several seconds is sufficient, but a real-time implementation would require a higher acquisition rate. We anticipate using 2D array US transducers for a real-time implementation. These transducers would provide acquisition rates of fifty to one thousand volumes per second [72, 73]. These 2D array transducers could also be fairly small and placed closer to the region of interest.

A third issue deals with the laser delivery system. As shown in our experimental setup, a laser would have to be fired at the organ in free space. This occurrence is unlikely in practical situations. We are developing a fiber delivery tool that will allow us to safely guide the laser beam into the patients body. This tool will also be able to project concurrent laser spots, greatly enhancing our registration acquisition rate. Another consideration is the number, size, and shape of the PA markers. In general, more PA markers are advantageous, because each marker provides additional information that one can use to either improve the tracking and registration result or to verify or evaluate an existing tracking and registration result. However, this also assumes that the PA markers can be uniquely distinguished from one another. The shape and size of each PA marker have similar advantages and disadvantages, so we will discuss them together. A larger size and more unique shape may improve the accuracy in camera segmentation procedures because they provide any segmentation solution with more features or information. However, the PA markers are not physical markers and are instead generated onto some unknown tissue surface. Thus, as these

spots become larger, the point source assumption begins to fall apart and the resulting shape may depend greatly on the surface structures.

The computation times shown throughout the methods and results section still require some optimization for this method to become real-time at a reasonable refresh rate. The most obvious areas for improvement are data collection and PA image beamforming. The data collection phase can be improved dramatically with two changes. First of all, automated data collection as opposed to manual data collection would theoretically bring the data collection to the laser system pulse rate, which in our case is 10Hz. The second issue would be the laser delivery system described above. Processing a single volume as opposed to a volume for each PA signal will greatly decrease computation time. Photoacoustic beamforming computational cost is similar to conventional and current B-mode beamforming methods. Therefore, we do not anticipate technical or computational challenges to implement real-time PA beamforming methods.

There are several of factors that will affect this systems errors as we move from a bench-top setup to *in vivo* experiments. When our SC system is replaced with a stereo endoscopic camera, the errors may increase because our SC system has a larger disparity due to the shorter distance between the two cameras in a stereo endoscopic camera. The disparity of a SC system directly affects the error in triangulating points found in each image into a 3D point. Further work will be done to quantify the effects of this change. Also, the errors were reported based on surface points. Since the region

of interest is often subsurface, our reported TRE will be biased for subsurface target errors. We believe that the bias will be fairly small since the PA spots are being detected in the same modality as any subsurface regions. Another factor is the effect of imaging a different medium using US. US images are generally reconstructed using a single speed of sound even though an image can contain multiple mediums with multiple speed of sounds. There is significant variance in the speed of sounds in the phantoms that we used as *ex vivo* tissue and gelatin have different speed of sounds. This heterogeneity affects the axial scaling of the US image, but this is a problem that any US application must deal with.

4.8 Conclusion

We have proposed an innovative 3D US to video direct registration medical tracking technology based on PA markers and demonstrated its feasibility on multiple *ex vivo* tissue phantom. In this paper, we showed the ability to generate a PA signal on multiple *ex vivo* tissue phantoms in various scenarios and PA spot localization errors rivaling point source localization errors found in SC systems. The TRE results have been shown to have higher accuracy than state of the art surgical navigation systems. Future work will include the development of a fiber delivery tool, spot finding algorithms to support concurrent spot projection, and subsequent *in vivo* animal experiments. Integration of this direct registration method into laparoscopic or robotic

surgery environments will also be a point of emphasis.

Chapter 5

Three-dimensional ultrasound to video registration with concurrent photoacoustic markers

5.1 Introduction

This chapter extends the work presented in chapter 4, and serves as another step towards realizing a practical clinical system shown in Figure 4.1B. The main contribution of this work is the development of a fiber delivery device, capable of projecting and generating concurrent PA markers. There are many advantages to using such a device in place of the prior methods of sequentially generating PA markers. Prior work used a free-space laser to generate the PA markers. There are safety concerns

with this approach as the laser light is allowed to freely move outside of the body cavity. It would also be difficult to create a complex system of mirrors that would guide the laser light into the body cavity. Thus, a fiber delivery device is a logical progression as it allows the user to easily manipulate the laser source. The output of the fiber can now be inside the body cavity, eliminating some of the safety concerns. The ability to concurrently project PA markers significantly decreases the acquisition time for this method. When the PA markers were being generated sequentially, an US volume must be acquired for each PA marker. With this concurrent PA marker fiber delivery device, only a single US volume is necessary.

This chapter details the design and development of the fiber delivery device, enabling concurrent PA markers, the extension to the software to support concurrent spots, and experimental results on a synthetic phantom with excellent light absorption characteristics and on an *ex vivo* porcine kidney embedded in a gelatin phantom.

5.2 Contributions

The main contributions of this chapter are the following:

- Extension of a photoacoustic-based tracking method for three-dimensional ultrasound transducer to use and process multiple photoacoustic markers simultaneously
- Experimentation on phantoms and *ex vivo* tissue to validate this extended track-

ing method

- Demonstration of target registration errors when using this tracking method that are at worst comparable with state of the art methods

5.3 Acknowledgement

I wish to thank Xiaoyu Guo for help with the optical setup and Hyun-Jae Kang for help with the MUSiiC toolkit. I was responsible for leading all components of this work.

5.4 Technical Approach

5.4.1 Fiber Delivery Design

The design of the fiber delivery device must meet several requirements. First of all, the input side must be fiber-coupled with the laser source. Standard optical fiber connectors can effectively couple the laser energy from the laser source into the optical fiber. These optical fiber connectors can be seen in Figure 5.1. The fiber must be able to carry and deliver enough energy to generate a PA marker on tissue. We will later show in our experiments that a bundle consisting of 200μm fibers is sufficient. The fiber must also be a fiber bundle since we want to have concurrent PA markers. This is only true because we use a single laser source. It is possible to use separate fibers if

there are multiple laser sources. To achieve concurrent PA markers, the fiber output must be able to project spots of light onto some surface 10-20cm away. This distance is chosen because we wish to eventually develop these tools for laparoscopic use. We separate the individual fibers in the fiber bundle and lightly focus the output using optical lenses, allowing each of these outputs to resemble spots on the target surface. Two different approaches were considered: independent lenses and single lens.

5.4.1.0.1 Independent Lenses

The approach of using independent lenses means that each of the individual fibers had an attached lens. This setup can be seen in Figure 5.1. The fiber is fixed at a certain distance away from a bi-convex lens. The fiber has a numerical aperture of 0.22. The bi-convex lens has an aperture of 6mm and a focal length of 10mm. The bi-convex lens lightly focuses the output and a single PA spot can be projected onto a surface 10-20cm away. A custom holder was created using a three-dimensional printer to hold the bi-convex lens stationary relative to the fiber output.

This design has several advantages and disadvantages. First of all, there is no set pattern to the concurrent PA markers. Each of the fiber outputs can move independent of each other. This can be seen as both an advantage and a disadvantage. As an advantage, no set pattern means that the pattern can be adjusted to best fit the region of interest. There may be regions of the surface that should be avoided. For example, if a PA marker is generated onto an edge between two types of tissue, this

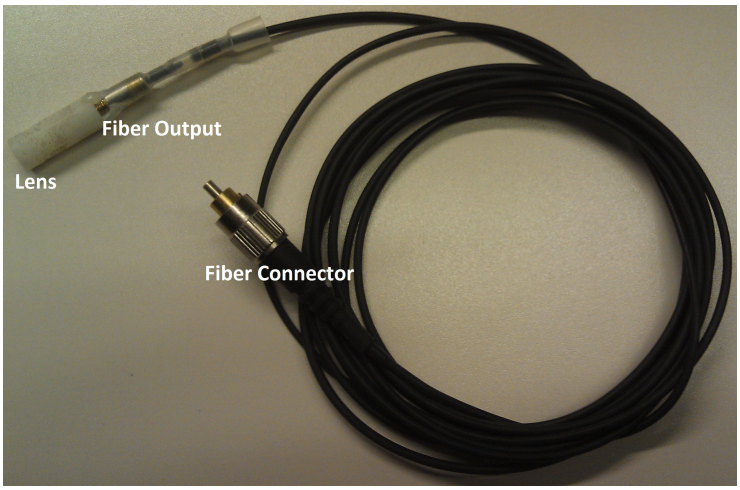


Figure 5.1: The fiber setup using independent collimating lenses (Far).

may possibly affect the quality of the PA marker seen in the US volume. However, the disadvantage is that there must be some method to reassemble the individual fibers into some pattern within the body cavity. This can be seen as added complexity, as these fibers must be somehow actuated and manipulated within the body cavity. Another advantage is that each individual fiber can be very small. If each fiber is introduced into a laparoscopic environment independently, it is likely that no additional port will be necessary. It may also be possible to attach each of these fibers onto another device or tool, such as the SC.

5.4.1.0.2 Single Lens

The alternative approach of using a single lens means that all of the fibers share a single lightly focusing lens. This setup can be seen in Figure 5.2. The fiber bundle is mechanically split by a fiber splitter into a set pattern. A convex lens is then attached

to the end of the tool. This convex lens has an aperture of 25mm and a focal length of 50mm. This device has also been designed to be adjustable. The fiber splitter can be adjusted to generate a pattern from a small pattern to a large pattern. The distance between the convex lens and the fiber output can also be adjusted to change the ideal distance for PA marker projection. A closeup of this adjustable device can be seen in Figure 5.3.

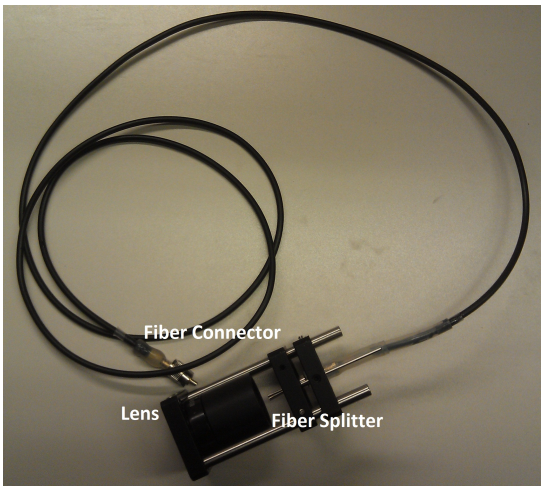


Figure 5.2: The fiber setup using a single collimating lens (Far).

This design also has several advantages and disadvantages. The projected pattern is relatively fixed, only differing in scale and skew parameters. This can act as additional information and aid the segmentation algorithms in finding all of the PA markers. A disadvantage of this approach is that the entire device is now a single large tool. This means that an additional port is necessary to introduce this tool into a laparoscopic environment. We decided to use this design in our experiments as its

easily adjustable nature allowed us more testing opportunities throughout the design process.

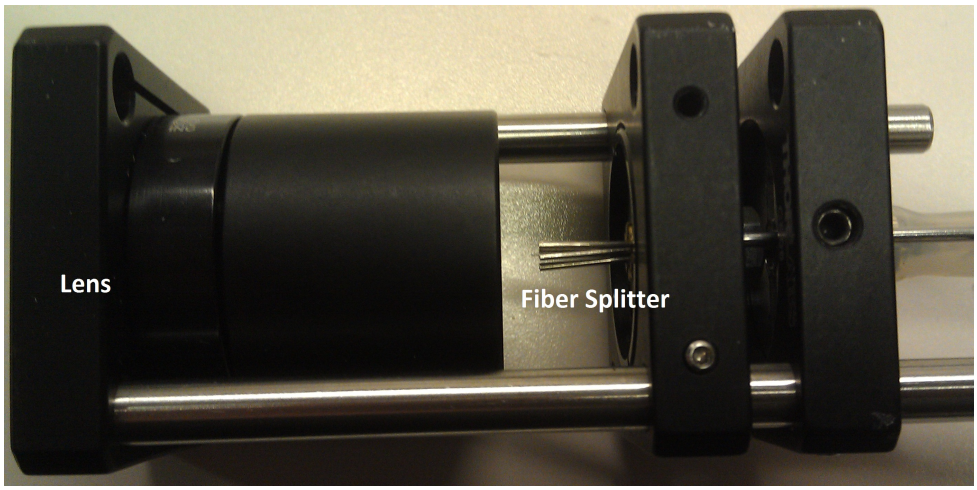


Figure 5.3: The fiber setup using a single collimating lens (Close).

5.4.2 Technical Experimental Workflow

The experiments can be separated into three phases: data collection, data processing, and registration. The data collection phase consists of collecting multiple pairs of SC images and a three-dimensional RF US volume. The data processing phase will then process the data to generate a three-dimensional SC point set and a three-dimensional US point set. These two point sets are registered together in the registration phase to finally output the transformation registering the SC frame to the US frame. The new workflows can be seen in figure 5.4.

The details of these phases are similar to those in figure 4.4 except for several

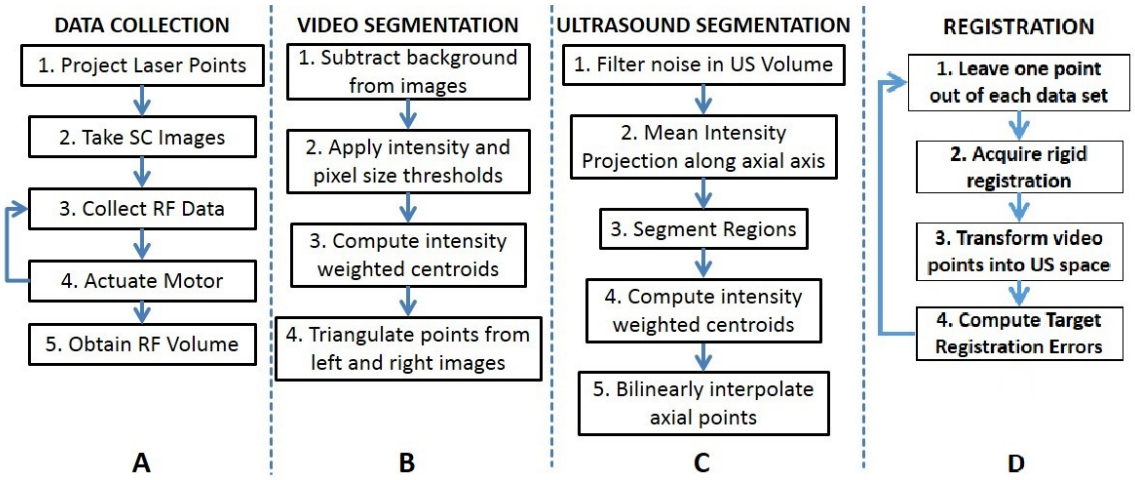


Figure 5.4: The workflows for the A) Data Collection, B) Video Segmentation, C) Ultrasound Segmentation, and D) Registration phases.

key differences. First of all, concurrent PA markers allows us to collect all of the necessary data in a single three-dimensional RF US volume and multiple pairs of SC images from the same instance. Previously the data collection pipeline was repeated for each PA marker, but that is no longer necessary. This significantly decreases the data acquisition time and makes the assumption that the three-dimensional RF US volume is at a single time instant much more plausible.

The other main change is modifying both the video and US segmentation methods to segment multiple markers as opposed to a single one. The same pipeline modified to accept more than a single PA marker worked fairly well. However, there were some cases where some of the PA markers could not be automatically segmented from the three-dimensional RF US volume. The energy density of each PA marker was non-uniform, making it difficult to automatically select a threshold that would be ideal for all of the PA markers. A more robust segmentation method not based solely on

intensity is in development, but this method is robust to missing PA markers because there is built-in redundancy in the PA markers.

The registration phase is unchanged from chapter 4. We note here that the coherent point drift [44] algorithm is used, so point correspondence is not necessary. This is a tremendous advantage for this method, as the PA markers seen in the SC frame and seen in the US frame will not have inherent correspondence since they are being projected concurrently. They will be corresponding point sets, but their correspondence will be unknown unless established with another method. It is still possible to use point set registration methods requiring correspondence if this point correspondence is first established.

5.5 Methods

In these experiments, we used a Q-switched neodymium-doped yttrium aluminum garnet (Nd:YAG) Brilliant (Quantel Laser, France) laser to generate the PA marker. We used a wavelength of 532nm and an energy density of approximately $1.6 mJ/cm^2$ on the synthetic phantom and approximately $3.8 mJ/cm^2$ on the *ex vivo* kidney phantom.These values are below the maximum permissible exposure (MPE), $19.5 mJ/cm^2$, as calculated from the IEC 60825-1 laser safety standard [68] based on a 0.25s exposure time, a 4ns pulse width, and a frequency of 10 Hz. We used a SonixCEP US system and a 4DL14-5/38 US transducer developed by Ultrasonix Medical Corpora-

tion (Richmond, Canada) to scan the volume of interest. This three-dimensional US transducer consists of a linear US array, motor actuated to move angularly around an internal pivot point. It has a bandwidth of 5 to 14MHz and the linear array is approximately 38.4mm. The Sonix DAQ device, developed by the University of Hong Kong and Ultrasonix, and the MUSiiC toolkit [52] are used to acquire prebeamformed radio-frequency (RF) data from the US machine. The k-wave toolbox [67] in MATLAB (Mathworks Inc. Natick, Massachusetts) is used to beamform and reconstruct PA images based on the prebeamformed RF data. The SC setup consists of two CMLN-13S2C cameras (Point Grey Research, Richmond, Canada) to capture images at 18Hz. The camera calibration process using the Camera Calibration Toolbox for MATLAB [69] generates a calibration file for the SC setup, allowing us to perform three-dimensional triangulation.

These experiments were performed on two phantoms. The synthetic phantom is made with plastisol and black dye. The *ex vivo* kidney phantom is made with gelatin and a freshly resected porcine kidney. The surface of the kidney is exposed from the gelatin such that there is an air-kidney interface. Figure 5.5 shows the *ex vivo* kidney phantom during the experiment. The PA markers within figure 5.5 has been artificially marked for presentation.

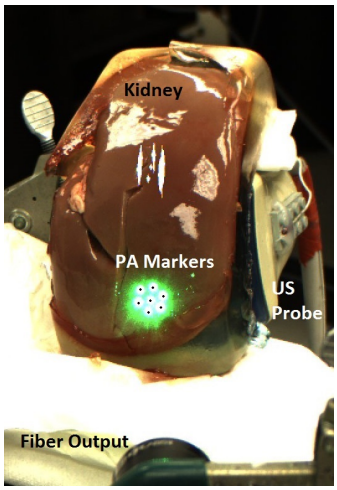


Figure 5.5: The experimental setup showing the PA markers generated on the kidney phantom surface.

5.6 Results

The registration results of our experiments on the synthetic phantom and the *ex vivo* kidney phantom are validated using the target registration error (TRE) metric defined in Equation 4.2. F_{SCUS} is the transformation between the SC frame and the US frame computed with all of the SC and US points except for one. The TRE is the difference between the actual US test point and the transformed SC test point in the US frame. N is the number of points in the experiment and N-1 points are used to compute F_{SCUS}. This computation is repeated with each of the N points as test points. The TRE results for the synthetic and *ex vivo* kidney phantom are shown in Table 5.1.

Phantom	Synthetic	*Ex vivo* Kidney Phantom
Lateral (mm)	0.38±0.28	0.48±0.45
Axial (mm)	0.8±0.10	0.46±0.34
Elevational (mm)	0.59±0.50	1.02±0.73
Euclidean Norm (mm)	0.84±0.37	1.36±0.64

Table 5.1: TRE results for the synthetic and *ex vivo* kidney phantom.

5.7 Discussion

The experimental results in Table 5.1 show that this three-dimensional US to video registration method using PA markers has higher accuracy than state of the art surgical navigation systems. While these are good results, there is some concern that they are worse than the results shown previously in section 4. There may be several explanations for this occurrence. First of all, the sample size of the point sets is very small. This is a result of using a single laser source to generate multiple PA markers. There is a limit to how many PA markers can be generated. There are several possible solutions. It is possible to project concurrent PA markers multiple times. This would maintain a short data acquisition time, while increasing the amount of data. Another solution is to use another laser source that can support more fibers. Second, the points are projected much closer together. This will cause any errors or uncertainties

in the point segmentation to be magnified because their magnitude will be fairly large relative to the distance between the points.

There are also some considerations in moving this system to *in vivo* experiments. Thus far, we have assumed that everything remains static during data acquisition. Evidently, this is not a valid assumption during an *in vivo* experiment. While we have decreased the acquisition time significantly, it is still on the order of 3-10s. This much time is required because the US transducer that we are currently using is an actuated transducer. Thus, we need to collect data for each actuated motor position. A two-dimensional array transducer would be able to provide an entire volume, thus reducing the acquisition time to a level where the assumption that the surgical environment is static is valid.

5.8 Conclusion

We demonstrated an extension to an innovative three-dimensional US-to-video direct registration medical tracking technology based on PA markers. We demonstrated the feasibility of this method using concurrent PA markers on a synthetic phantom and an *ex vivo* kidney phantom. We showed that this method has higher accuracy than state of the art surgical navigation systems. Future work will explore improving the robustness of the segmentation algorithms, *in vivo* animal experiments, and integration of this registration method into laparoscopic or robotic surgery environments.

Chapter 6

Two-dimensional curvilinear ultrasound to video registration with photoacoustic markers

The main contribution of this chapter is a method to recover the video to ultrasound registration with two-dimensional images from a single pose. Naturally, this means that this method can be extended to two-dimensional US transducers and is no longer limited to three-dimensional US probes. There are both advantages and disadvantages of using two-dimensional US transducers versus three-dimensional US transducers. Three-dimensional transducers are capable of providing an entire volume at the expense of acquisition time. On the other hand, two-dimensional transducers are much faster, but the registration between two-dimensional US images and three-

dimensional video is considerably more difficult. Using a two-dimensional transducer will also allow us to relax the assumption that the surgical environment remains static during data acquisition as a volume is no longer necessary.

In this chapter, we show that it is possible to recover some out-of-plane information from a single two-dimensional curvilinear US image when using PA markers. We detail the ideas and algorithms that facilitate the process of registering two-dimensional US images with three-dimensional video. Preliminary results using a synthetic phantom with excellent light absorption characteristics will also be shown.

6.1 Contributions

The main contributions of this chapter are the following:

- Extension of a photoacoustic-based tracking method for use with two-dimensional ultrasound transducers
- Experimentation on phantoms to validate this extended tracking method
- Demonstration of target registration errors when using this tracking method that are comparable with state of the art methods

6.2 Acknowledgement

I wish to thank Xiaoyu Guo for help with the optical setup and Hyun-Jae Kang for help with the MUSiiC toolkit. I was responsible for leading all components of this

work.

6.3 Technical Approach

There are two main components that allow us to recover the registration between video and ultrasound with a single US image. First of all, a wavefront segmentation algorithm is necessary as the wavefront allows us to recover some out-of-plane information. This wavefront represents the time of flight (ToF) readings from the active PA marker to each of the US transducer elements. Second of all, to be able to register PA markers between two-dimensional US and video, we show how the ToF readings can be used to recover out-of-plane information. Before we present these two components, we will discuss the model that we use to derive the subsequent algorithms.

6.3.1 Wavefront Modeling

As we previously mentioned, the wavefront from a single PA marker can represent the ToF between the PA marker and each transducer element. As shown in figure 6.1, we have a set of transducer elements from a curvilinear transducer, e_i, and a single PA marker, p. R represents the transducer curvature and θ represents the element pitch. The set of distances between each element e_i and p is represented by w_i. As we can see from this model, we are making the assumption that each of the transducer

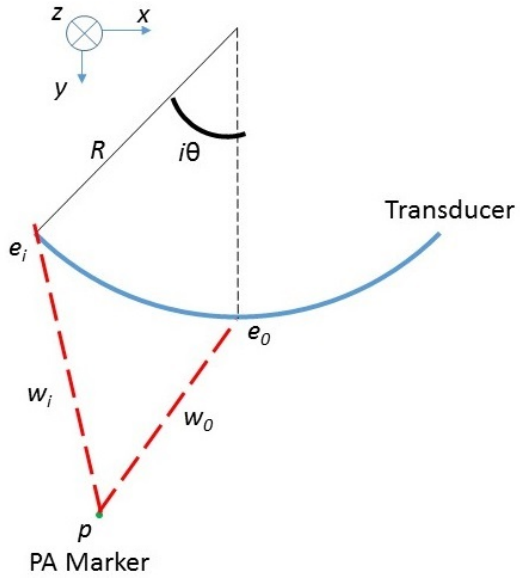


Figure 6.1: Concept figure and definition of variables in wavefront modeling.

elements and the PA marker itself are acting as point receivers and a point source respectively.

6.3.2 Wavefront Segmentation

Segmentation of a point source in a PA image is generally much easier than in an US B-mode image. In most cases, an algorithm using intensity thresholds is sufficient in segmenting the desired wavefront. One can then look at the radio-frequency (RF) signal for each element and pick the earliest signal as the wavefront's ToF for that particular element. Naturally, this will only work when there is a single wavefront present in the PA image. When there are wavefronts from multiple PA markers present in a single PA image as seen in figure 6.2, we can no longer just pick the

earliest signal as the wavefront's ToF.

$$w_{ij}^2 = a_j + cos(i\theta)b_j + sin(i\theta)c_j \tag{6.1}$$

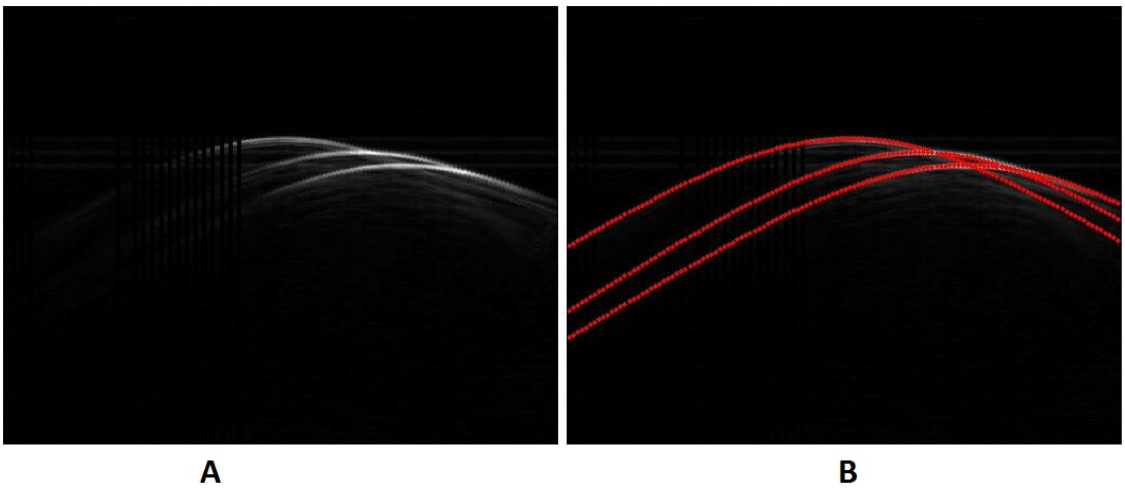


Figure 6.2: Sample pre-beamformed PA image with PA markers present. A) Without Segmentation B) With Segmentation

Based on the wavefront model shown in figure 6.1, each wavefront from a unique PA marker can be represented by equation 6.1. There is a slight change of notation where w_{ij} is the axial position of the wavefront j for element i. In addition, a_j, b_j, and c_j are constants unique to PA marker j. From the initial intensity thresholding algorithm, we acquire a set of w_i with unknown correspondence to a set of PA markers. This means that there may be multiple w_i values for any particular element i. We combine equation 6.1 with a RANSAC approach [74] to separate these w_i into w_{ij}. A set of w_i for different is are selected randomly. We then solve for a, b, and c, which will uniquely define the equation for a wavefront in the PA image. We then find all w_i that satisfy this wavefront and determine if there are enough votes from w_i to state

that this wavefront corresponds with a unique PA marker j. If so, this set of w_i is removed from the complete set and denoted as w_{ij}. The derivation for equation 6.1 will be presented in the next section.

6.3.3 Out-of-plane Position Recovery

The segmented wavefronts can be used to recover some information about the PA marker's position relative to the US transducer. As was previously mentioned, each w_i represents the ToF and corresponding distance between element i and the PA marker's three-dimensional position. If we consider the ideal case, where the transducer elements are ideal point sources, we can represent this relationship with equation 6.2 where e_i is the position of element i and p is the position of the PA marker. Since we are using curvilinear transducers, e_i can be described by equation 6.3 where R is the radius of the curvilinear transducer. Substituting e_i into equation 6.2 will result in a form that satisfies equation 6.1. Theoretically, p_i can be uniquely solved up to sign in this situation given a minimum of 3 valid w_i. There are several approaches to solve this equation. One approach is to directly solve equation 6.2 with a non-linear optimization approach. Another approach is to use the unique constants a, b, and c, from equation 6.1 and solve for the p terms within those constants. The unknown sign means that the point can either be in front or behind of the image plane.

$$w_i^2 = (e_{ix} - p_x)^2 + (e_{iy} - p_y)^2 + (e_{iz} - p_z)^2 \tag{6.2}$$

$$e_i = [Rsin(i\theta), Rcos(i\theta), 0] \tag{6.3}$$

6.3.4 Technical Workflow

The experiments can be separated into three phases: data collection, data processing, and registration. The data collection phase consists of collecting a pair of SC images and a two-dimensional prebeamformed PA image. The data processing phase will then process the PA image using the aforementioned wavefront segmentation and out-of-plane position recovery algorithms to generate two p_i for each PA marker. The PA markers are segmented from the SC images using the same methods described in sections 4 and 6. These two sets of information are registered together in the registration phase to finally output the transformation registering the SC frame to the US frame. The new workflows can be seen in figure 6.3. This work mainly focused on the wavefront segmentation and PA marker position recovery under ultrasound segmentation.

The key difference in these phases between this work and previous work is the registration phase. Since each PA marker in the US coordinate system can only be solved up to a sign, there are now a pair of possible points for each PA marker. We use all of these points in the US coordinate system and register them with the PA markers in the SC coordinate system using the coherent point drift [44] algorithm.

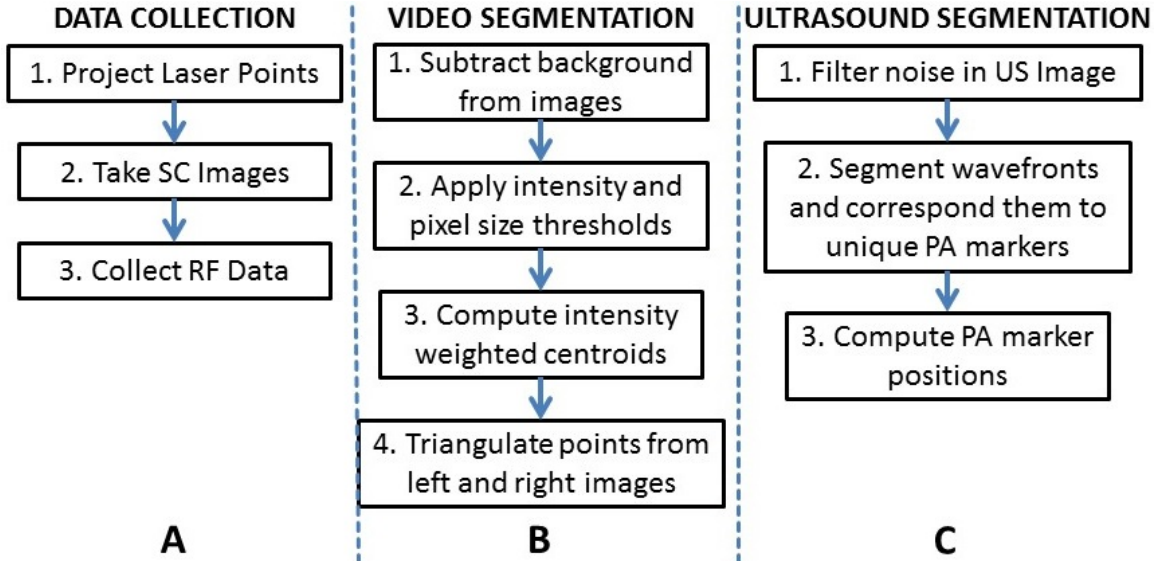


Figure 6.3: The workflows for the A) Data Collection, B) Video Segmentation, and C) Ultrasound Segmentation.

This algorithm is ideal for these cases where point correspondence is unknown and the number of points in each dataset are different. The one thing that must be noted is that we must check that the resulting transformation is in SE(3) and is not a reflection.

6.4 Methods

In these experiments, we used a Q-switched neodymium-doped yttrium aluminum garnet (Nd:YAG) Brilliant (Quantel Laser, France) laser to generate the PA marker. We used a wavelength of 1064nm and an energy density between 5 to $10 mJ/cm^2$ on the synthetic phantom.These values are below the maximum permissible exposure (MPE), $100 mJ/cm^2$, as calculated from the IEC 60825-1 laser safety standard [68] based on

a 0.25s exposure time, a 4ns pulse width, and a frequency of 10 Hz. We used a Sonix Touch US system and a 4DC3-7/40 US transducer developed by Ultrasonix Medical Corporation (Richmond, Canada) to collect two-dimensional images. This three-dimensional US transducer consists of a curvilinear US array, with a motor actuated to move angularly around an internal pivot point. For the purposes of this work, we used it as a two-dimensional transducer and did not actuate the motor. It has a bandwidth of 3 to 7MHz and the transducer radius is approximately 40mm. The Sonix DAQ device, developed by the University of Hong Kong and Ultrasonix, and the MUSiiC toolkit [52] are used to acquire prebeamformed radio-frequency (RF) data from the US machine. The SC setup consists of two CMLN-13S2C cameras (Point Grey Research, Richmond, Canada) to capture images at 18Hz. The camera calibration process using the Camera Calibration Toolbox for MATLAB [69] generates a calibration file for the SC setup, allowing us to perform three-dimensional triangulation. These experiments were performed on a synthetic black plastisol phantom.

6.5 Results

The registration results of our experiments on the synthetic phantom are validated using the target registration error (TRE) metric defined in equation 4.2. F_{SCUS} is the transformation between the SC frame and the US frame computed with all of the SC and US points except for one. The TRE is the difference between the actual US test

point and the transformed SC test point in the US frame. N is the number of points in the experiment and N-1 points are used to compute F_{SCUS}. This computation is repeated with each of the N points, 6 in this case, as test points. The resulting TRE for this experiment was 2.47mm and the standard deviations was 1.29mm, which is comparable to current available systems.

6.6 Discussion

The experimental results show that this three-dimensional US to video registration method using PA markers from a single image has comparable accuracy to state of the art surgical navigation systems. While there is some concern that they are worse than the results shown previously in sections 4 and 6, it is also expected as there is much less information present in a single PA image as opposed to an entire volume. Besides this reason, there may be several explanations for this occurrence. Errors in both the wavefront segmentation and out-of-plane position recovery algorithms will naturally propagate to the PA marker positions. Since the position is recovered based on ToF, the segmented wavefront is ideally the earliest arrival of acoustic pressure. However, what ends up being detected may be the earliest arrival of acoustic pressure that is greater than noise. Also, some speed of sound must be used to convert the ToF for out-of-plane position recovery. The largest source of error likely stems from the initial assumptions that were made. In this work, we assume that the US transducer

consists of ideal point source receivers. Future work is aimed at devising a more suitable model without making such strong assumptions.

There are also some considerations in moving this system to *in vivo* experiments. One advantage of using two-dimensional US images versus a three-dimensional US volume is that the data-acquisition time is shortened drastically. Previously, the assumption that the environment is static while the volume is being collected had to be made. However, with a registration method that uses a single US image, we no longer have to make this assumption. This greatly decreases the barrier between this method and *in vivo* experiments.

6.7 Conclusion

We demonstrated an extension to an innovative three-dimensional US-to-video direct registration medical tracking technology based on PA markers using US images from a single pose. We demonstrated the feasibility of this method on a synthetic phantom. We showed that this method has comparable accuracy with state of the art surgical navigation systems. Future work will devise a more realistic model, quantify the accuracy of recovering out-of-plane PA markers, *ex vivo* experiments, *in vivo* animal experiments, and integration of this registration method into laparoscopic or robotic surgery environments.

Chapter 7

Bi-planar trans-rectal ultrasound to video registration with photoacoustic markers

7.1 Introduction

In chapter 6, we demonstrated the feasibility of using a convex array transducer in place of a three-dimensional transducer. While this change improved the practicality of the system by reducing the length of time and the static environment assumption, the accuracy also decreased because of the limitations in the physical arrangement of the convex array transducer elements. In a convex array, the elements lie on a circular arc. The constraints of this geometry is lacking because the elements lie on

a two-dimensional shape.

In this chapter, we show how a bi-plane transrectal transducer, with its elements lying in a three-dimensional configuration, can improve the accuracy, while maintaining the advantage of not using a three-dimensional transducer. We detail the key idea that enables our method and demonstrate experimental results in two scenarios with different active point sources.

7.2 Contributions

The main contributions of this chapter are the following:

- Extension of a photoacoustic-based tracking method for use with bi-plane transrectal ultrasound transducers
- Experimentation on phantoms to validate this extended tracking method
- Demonstration of target registration errors when using this tracking method that are comparable with state of the art methods

7.3 Acknowledgement

I wish to thank Younsu Kim for help with the PZT setup. I was responsible for leading all components of this work.

7.4 Technical Approach

7.4.1 Video to Ultrasound Registration with Photoacoustic Markers

As can be seen in figure 7.1, the workflow for this method remains very similar to the ones shown in sections 4, 5, and 6, so we will briefly summarize it here. For each photoacoustic marker projection, a pair of stereo camera images are taken. Simultaneously, the channel data from the transducer, in this case a transrectal biplane transducer, is also acquired. The stereo camera pair is used to recover the position of the photoacoustic marker in the stereo camera systems coordinate system. The channel data is used to recover the position of the photoacoustic marker in the ultrasound transducers coordinate system. The process of recovering this position is different than in previous work and will be described in the following section. After at least three photoacoustic markers, a registration can be found between the two point sets using any standard point set registration algorithms such as coherent point drift [44]. This registration is equivalent to tracking the ultrasound transducer in the stereo camera systems coordinate system.

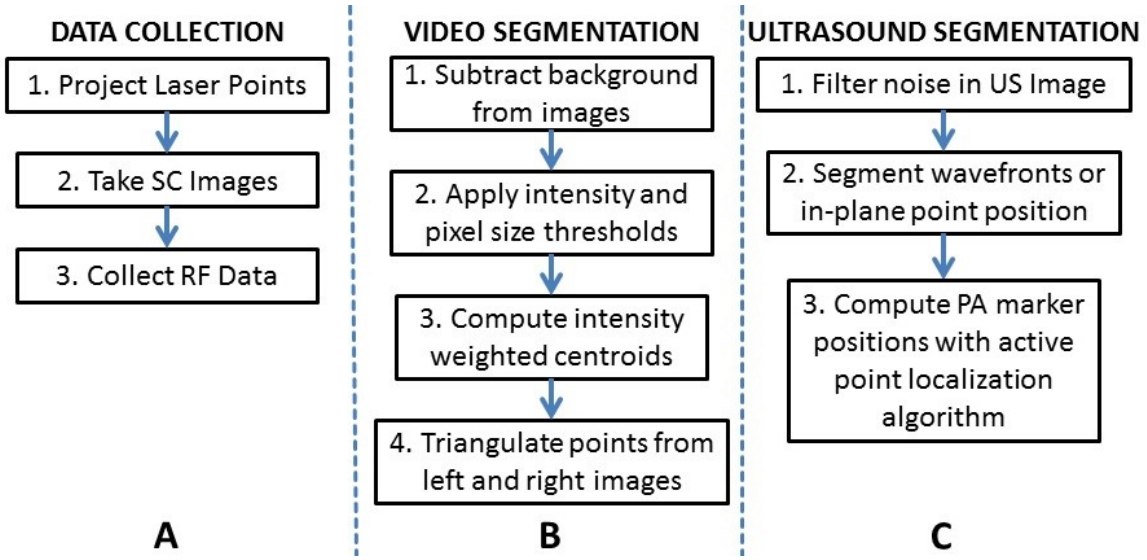


Figure 7.1: Experimental and Software Workflow

7.4.2 Active Point Localization with a Transrectal Transducer

Before we begin describing the method for localizing the active acoustic point source with respect to the transrectal transducer, we must describe the transducer itself. Bi-plane transrectal transducers typically have two imaging planes, one parallel and one perpendicular to its insertion axis. The parallel imaging plane is generally from a linear array and the perpendicular imaging plane is generally from a convex array. Figure 7.2 is an example of such a transducer, where the dotted lines correspond to the imaging planes. In general, these two imaging planes are perpendicular to each other.

The key idea that enables our approach is the use of an active acoustic point

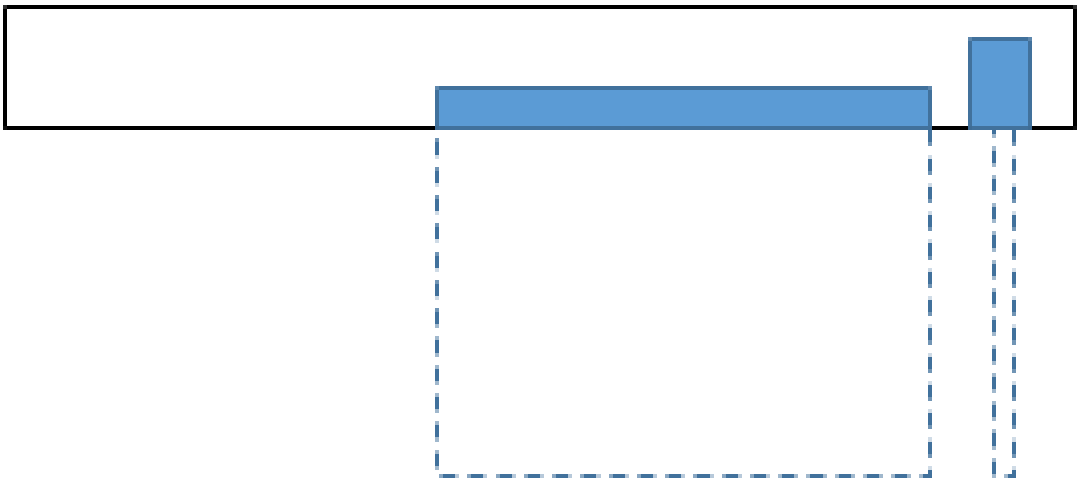


Figure 7.2: Ultrasound imaging planes with respect to bi-plane transrectal transducer.

source. This simply means an acoustic source where we have full control over its transmission and timing. When using an active point, the receiver will be able to capture data even if the point is outside of the probe imaging plane. Figure 7.3 shows an example of what the prebeamformed data of an active acoustic point source might look like in a bi-plane transrectal transducer. As highlighted by the line, it can be clearly seen that there is a delineation between the data acquired from the convex array versus the linear array.

For example, in the scenario shown in figure 7.4 where the point represents an active acoustic point source, both of the transducers will receive the signal from the active acoustic point source. This property significantly increases the field of view of an ultrasound transducer. Naturally, there are still some constraints related to the transducers receiving specifications. For example, it is unreasonable to expect the signal to be received if the active acoustic point source is tens of centimeters away

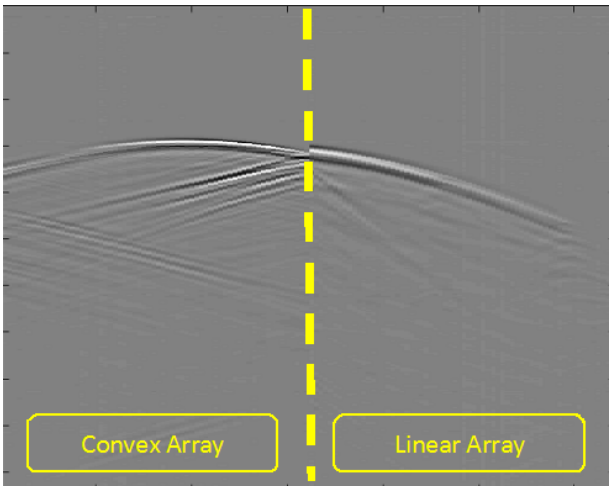


Figure 7.3: Sample pre-beamformed channel data from a bi-plane transrectal transducer.

normal to the imaging plane as there is some angular sensitivity to the transducers receiving elements. The actual constraint will depend on this angular sensitivity as well as the strength of the transmitted acoustic signal.

In this work, we present two main methods for localizing the active acoustic point source based on the data received by the bi-plane transrectal transducer. The main distinction between these two methods is that one uses requires pre-beamformed channel data, while the other uses beamformed data directly. The focus of our work is on the method using beamformed data, but the general method using pre-beamformed channel data is also presented for completeness.

We will first describe the geometrical model of an active point source. The situation when the ultrasound transducer is linear is shown in figure 7.5. The PA spot or any active point source is denoted as p and transducer element i is denoted as e_i. w_i represents the distance from the PA spot to transducer element i. Since we are

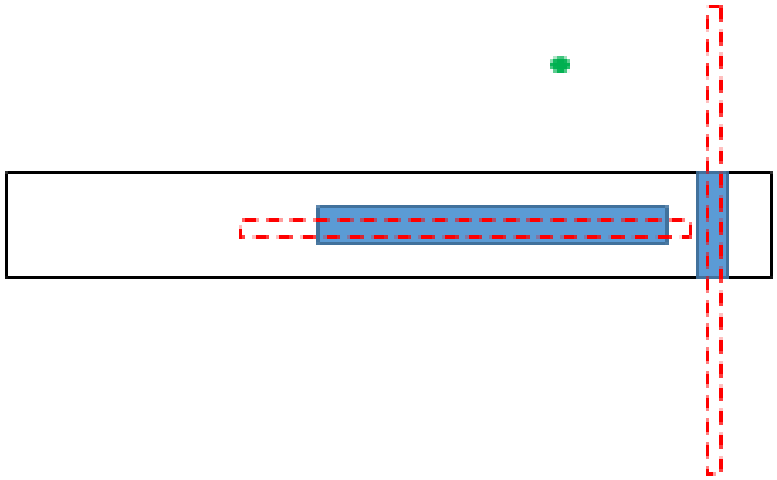
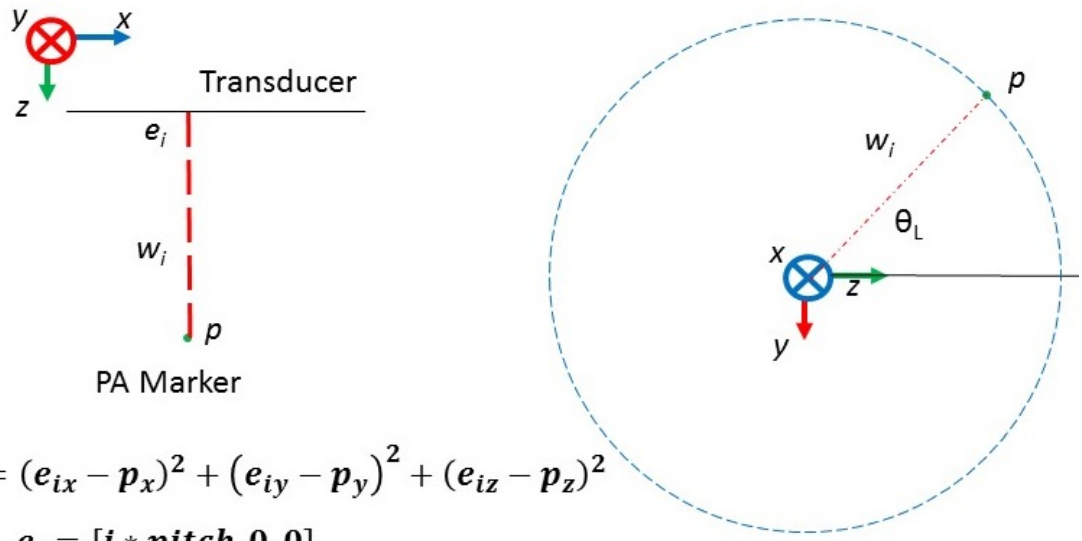


Figure 7.4: Scenario where active point is outside of the imaging planes.

looking at active points, its position in the beamformed image must be interpreted in a different manner than a typical pulse-echo ultrasound image. While we can still trust the lateral position of the point in the image, the axial position now contains an elevational component as well. From the beamformed image, we know the distance, w, from the PA spot to the closest element on the transducer, e_i. With just this information, there is one unknown degree of freedom, θ_L, representing a circle about element e_i that the PA spot can lie on. A similar model can be shown for a curvilinear ultrasound transducer, with the main difference being that e_i having some axial component, when using beamformed images only.

As this active point source is observed by both the linear and curvilinear array, there will be two such circles, each with one unknown degree of freedom. Ideally, the active point will lie on the intersection of these circles. Since the circles may



$$w_i^2 = (e_{ix} - p_x)^2 + \left(e_{iy} - p_y\right)^2 + (e_{iz} - p_z)^2$$

$$e_i = [i * pitch, 0, 0]$$

$$w_i^2 = (e_{ix} - p_x)^2 + p_y^2 + p_z^2$$

Figure 7.5: Geometrical model of an active point source observed by a linear ultrasound transducer.

not actually intersect, due to small calibration errors, we can solve this system in a least-squares sense to find the point coordinates. An alternative method (used in the experimental results below) computes the intersection of the circular arc determined from the curvilinear transducer with the plane containing the circular arc determined from the linear transducer image. The information provided by the linear transducer allows one to define the lateral position of the point. This fixed lateral dimension can be represented as a plane parallel to the curvilinear transducer. Thus, the circle of possible positions from the curvilinear data will intersect with this plane. In either case, solving the system also requires that we know the transformation between to two transducer coordinate systems, which can either be determined by calibration or obtained from the bi-plane transrectal transducer specifications.

There also exists a more general method for solving for the position of the active point source. This method uses the acquired channel data before beamforming is applied. Referring back to figure 7.3, each channel will have a signal received from the active point source corresponding to its time of flight (TOF). This same scenario can be seen in figure 7.4. Previously, we only used the w_i corresponding to the shortest TOF, but now we use the entire array of w. Given this information, one can set up an optimization problem of the form shown in equation 7.1. In this equation, e_i represents the position of each of the transducer elements, p represents the position of the active point source, and w represents the distance between each e_i and p. This type of method is only suitable for transducers with non-collinear elements like bi-plane transrectal transducers or curvilinear transducer, but not for linear transducers.

$$\forall i = 1 \dots n : \arg\min_{p \in R^3} (||e_i - p||_2 - w_i)^2 \tag{7.1}$$

7.5 Methods

We designed two experiments to test the feasibility of the proposed method. The first experiment focuses on the active point localization algorithm. To isolate this method from other effects such as stereo camera tracking errors, we introduce the use of an active piezoelectric (PZT) element as a replacement for the PA marker. While the active signal generation of a PZT element is similar to a PA marker, we must use

another way to track the PZT element as we want to avoid using a stereo camera. In this case, we chose to use a translational stage that placed the PZT element sequentially in a grid-like pattern. The known positions in this grid becomes a replacement for the PA markers tracked by the stereo camera. This registration provides the pose of the ultrasound transducer relative to the grid-like pattern, and not relative to some external tracking frame. The ultrasound channel data is collected using a SonixTouch ultrasound system, SonixDAQ data acquisition system, and BPC8-4/10 and BPL9-5/55 bi-plane transrectal transducer.

The second experiment incorporates photoacoustic markers and a stereo camera system, accomplishing the originally stated goal of tracking the ultrasound transducer. In this experiment, PA markers are generated sequentially, with stereo camera images and ultrasound channel data being acquired simultaneously. The photoacoustic setup consists of a Q-switched Nd:YAG laser and corresponding optical mirrors and lenses. The stereo camera system consists of two calibrated CMLN-13S2C cameras. Points from the 10Hz Nd:YAG laser are sequentially projected onto a black plastisol phantom and collected using the same ultrasound transducer and SonixDAQ as the first experiment. The energy from these spots are well below the IEC laser safety limits [68]. The software used to synchronize and acquire data from each of these data sources is based on the MUSiiC Toolkit [52].

7.6 Results

The first experiment (with an active PZT element) had a total of 16 points, resulting in target registration errors of 1.07mm±0.35mm with a leave one out analysis. The equation for this metric is found in equation 4.2. F is the computed transformation of the ultrasound transducer in the stereo cameras coordinate system, while SC and US are the left out photoacoustic marker in the stereo camera and ultrasound coordinate systems respectively. TRE is the mean of each test point (active PZT source or PA marker) being used as the test point iteratively. Naturally, the first experiment did not use the stereo camera and all of the active PZT points are defined with respect to its grid-like pattern. The second experiment (with PA markers) had a total of 7 points, resulting in target registration errors of 1.27mm±0.47mm with a leave one out analysis. Figure 7.6 is a graphical representation of the registration result of the first experiment.

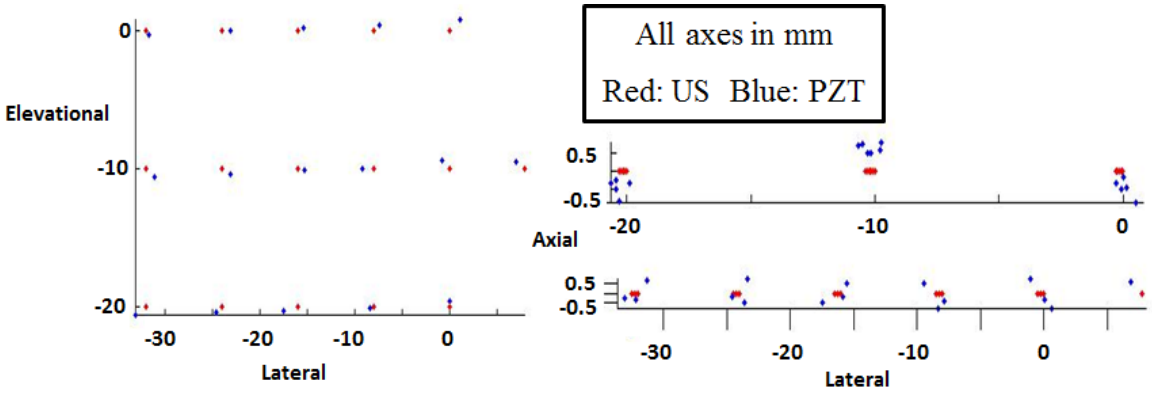


Figure 7.6: Result of point set registration for the first experiment.

7.7 Discussion

From these two experiments, we can see that the active point localization algorithm and video to bi-plane transrectal ultrasound registration using photoacoustic markers is feasible. The first experiment shows that the localization method is accurate to approximately 1mm, which is significant, given that the PZT element itself is approximately 2mm in diameter. From figure 7.6, we can also see that there seems to be some bias in the localization method. One possible reason for this is that the model is based on ideal point sources and point receivers. Relaxing this assumption will naturally affect the model, but will also more closely resemble the physical apparatus.

From the second experiment, we can see that the errors increase when including the stereo camera and photoacoustic systems. There are two main possibilities for this result. First, stereo camera calibration errors are typically sub-millimeter. This is negligible in most applications, but fairly significant in this case as our errors are already fairly low. The second possibility is that the active point is much larger in the photoacoustic case. The laser spots were roughly 6mm while the active PZT element was 2mm. This can lead to error in both stereo camera segmentation and active point localization as the point source is now much larger. A more general explanation for part of the error is a certain assumption that we have to make regarding the speed of sound in the medium. In our model, we assume a homogeneous medium with a known speed of sound. It is difficult in practice to accurately pick a speed of sound

for a certain medium. This is another possible cause for errors.

There are still several necessary advances to enable use of this technology in a practical setting. First of all, the sequential firing and generation of PA spots will take much too long in its current setup for real-time tracking. One obvious solution to this is to use a laser with a higher pulse repetition rate. Another solution is to use concurrent PA markers as we have previously shown in chapter 5. One reason why this was not the first choice is because the bi-plane transrectal probe cannot really be regarded as a single transducer. It acts more similarly to two transducers that are active at the same time. As such, if multiple signals are seen by each of the transducers, then there will be ambiguity in how we find correspondence between a single PA marker seen in each transducer. A model-based approach similar to the one we previously developed in chapter 6 may work, but that remains to be tested in the future.

7.8 Conclusion

In this work, we showed the feasibility of registering video and ultrasound using a bi-plane transrectal transducer with photoacoustic markers. This approach obtained results better than conventional surgical navigation systems. Future work will include concurrent photoacoustic marker projection and *ex vivo* experiments.

Part III

Interventional tool tracking with photoacoustic sources and single element ultrasound receivers

This part presents the work done related to tracking tools such as needles and catheters. The main objective is to integrate an acoustic sensor onto the tools themselves and track this sensor relative to some external tracking system such as optical cameras or ultrasound transducers. The following chapters explore different components related to this problem. Chapter 8 discusses a needle tracking method that makes use of a camera and an ultrasound transducer. Chapter 9 looks at an alternate pose tracking method itself with a marker made of light that we call the virtual rigid body. Chapter 10 explores a photoacoustic-based catheter tracking method that makes use of some of the techniques presented in part II.

Chapter 8

Needle tracking with mono-camera and piezoelectric sensor fusion

Needles are used in many surgical procedures such as drug delivery or needle biopsies. One of the key challenges when using needles in these interventions is the placement of the needle. Placement of the needle at the goal position will ensure proper execution of the surgical plan as well as avoid possible complications.

The tracking of needles, will generally make use of external tracking sensors such as optical tracking or electromagnetic (EM) sensing [18] to provide real time spatial information of the tool relative to the patient. Optical tracking systems require line of sight, while EM-based systems are wired and subjet to EM field distortions, discouraging the use of metallic tools. In addition, the estimation of the tool tips is limited by tool shaft bending and the effects of angle estimation error if the sensors

themselves are placed far away from the tip. Further, IOUS to camera or IOUS to tool tracking transformations necessarily require an indirect calculation based on a chain of spatial transformations, each with errors that may propagate to the next. The tools themselves are also often difficult to visualize within an IOUS image. Stoll et al. [19] attached passive markers on the surgical instrument such that its position and orientation could be determined from an ultrasound image. Rohling et al. explored image processing [20] and beamforming approaches [21] to enhance tool visibility.

This chapter explores the use of a needle with a piezoelectric sensor embedded at its tip. Similar to its use in part I, this can aid in visualization and detection of the needle tip within an ultrasound image. This sensor is insufficient for three-dimensional tracking by itself. Thus, a mono-camera is attached to the ultrasound transducer. On its own, it is also insufficient, but we show in this chapter how to combine these two sources of information to enable three-dimensional tracking of the needle tip.

8.1 Contributions

The main contributions of this chapter are the following:

- Development of a needle tracking method by fusing incomplete camera and ultrasound sensor information
- Experimentation to validate this extended tracking method

- Demonstration of sub-millimeter needle errors when using this tracking method in an ideal environment

8.2 Acknowledgement

This work was done in close collaboration with Bofeng Zhang and Philip Oh. I was a mentor to them when they worked on this project. I was involved in the algorithmic design, experimental design, and validation.

8.3 Technical Approach

8.3.1 Mono-camera needle segmentation and processing

Needle segmentation from camera images is required to obtain half of the necessary information to localize the needle tip. The needle appears as a line in the camera image. Since we are working with a single camera, the location of this needle is underdetermined. The three-dimensional physical location of each image point observed by a single camera can be modeled as a line extending from the camera's optical center through this point in the image with depth uncertainty. Thus, if we extend this to every point on the needle, we end up with a plane spanning each of these lines on

which the needle must lie on in three-dimensional space.

To segment the needle, we use a three step approach. First, we apply an intensity filter across the image to reduce the background of the image. Then, we apply a hough transform [75] to determine the locations of lines within the image. Finally, we apply a line length filter to only keep the longest, most well-defined set of lines. These set of lines typically include either edge of the needle and can then be averaged together to obtain the needle centerline.

$$N = (p_1 - o) \times (p_2 - o) \tag{8.1a}$$

$$v = p_1 - o \tag{8.1b}$$

We can then determine the plane on which the needle lies on by picking any two points, p_1 and p_2, on the segmented needle centerline. Referring to equation 8.1b where o is the camera's optical center, we can define the plane by its normal, N, and vector, v. This plane will be used later when we fuse it with the ultrasound information to obtain the needle tip position.

8.3.2 Ultrasound signal segmentation and processing

The piezoelectric element acts as an active acoustic source and we model it as an ideal point source. One important note is that active point sources can be observed

in the ultrasound image even if it is outside of the imaging plane. Active point sources are generally straight-forward to segment from ultrasound images because they either have higher intensity than the background or the data acquisition system can be configured such that there is no background at all. In this case, we use an intensity filter to determine the location of the active point as observed in the ultrasound image.

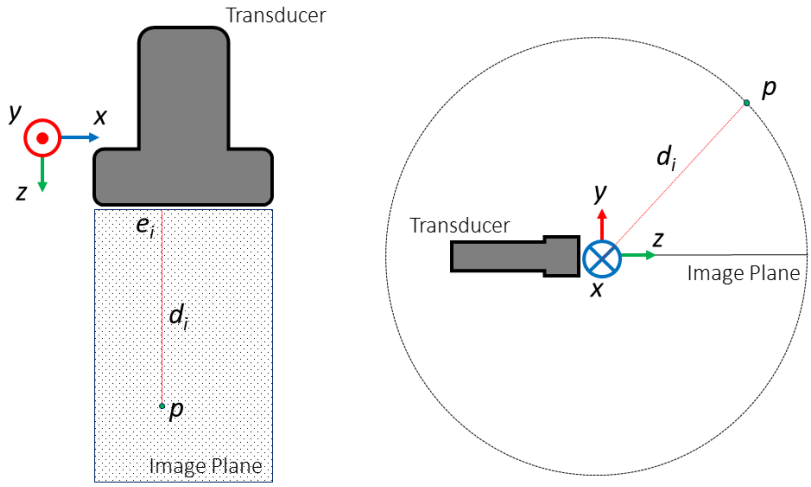


Figure 8.1: The concept for performing out-of-plane estimation. Given the lateral coordinate and the distance between the point and the transducer element closest to it, the point must exist on a circle within the axial-elevational plane.

We use the same out-of-plane estimation described in Chapter 3. Figure 3.5 is repeated here for the reader's convenience. This circular arc, C, can be parametrized as shown in equation 8.2, which will allow it to be easily transformed to another coordinate system. In this equation, e_i refers to the lateral position of the segmented piezoelectric signal and d_i refers to its axial position. t is then the parametrized angle defining the rotation of this point about the ultrasound transducer's lateral axis.

$$\forall t : C = \begin{bmatrix} e_i \\ d_i * sin(t) \\ d_i * cos(t) \end{bmatrix} \tag{8.2}$$

8.3.3 Camera and ultrasound sensor fusion

The information obtained by each of the individual sensors can only incompletely determine the position of the needle. However, they compensate for each other when combined together. The first step to combining the information from these two sensors is to put them in the same coordinate system. The plane is originally defined with respect to the camera, while the arc is defined with respect to the ultrasound transducer. Thus, one way of relating these two coordinate systems is through an ultrasound calibration process. A possible method was described previously in Chapter 3.

After the information from these two coordinate systems are calibrated such that they are within the same coordinate system, one can then fuse the information together to localize the position of the needle tip. Since the information from each sensor constrains the location of the needle tip to some set of possible locations, their intersection will result in the needle tip's location. Equation 8.3 describes this relationship analytically, where X is the calibration transforming every ultrasound point into the camera frame. This relationship is only satisfied when $X * C(t)$ lies on the camera and needle plane. Figure 8.2 is the graphical representation of this same

relationship.

$$(X * C(t) - o) \cdot N = 0 \tag{8.3}$$

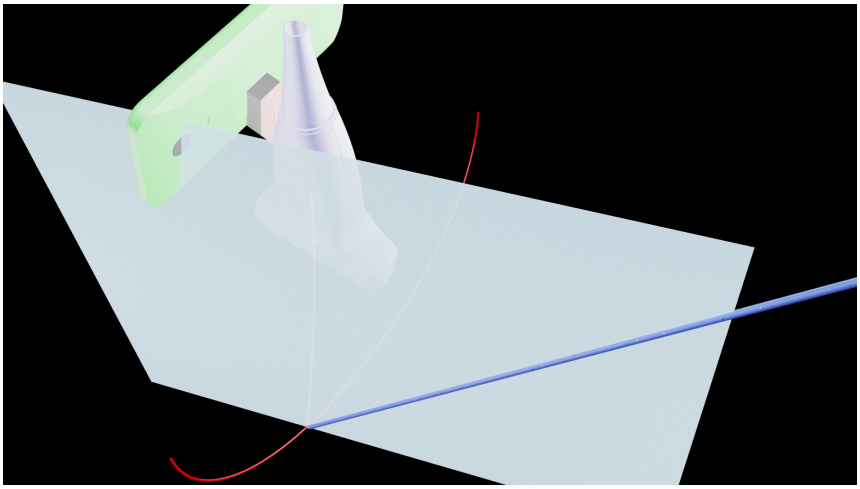


Figure 8.2: The graphical description of fusing the sensor information from the camera and piezoelectric element. The plane obtained from the camera and the arc obtained from the ultrasound are geometrically intersected to determine the location of the needle tip.

8.4 Methods

8.4.1 Apparatus

This needle tracking approach requires three main pieces of apparatus. The first is a camera to observe the shaft of the needle. In our experiments, we used a Motorola Moto G cellphone as the camera. The second is an ultrasound transducer and its associated data acquisition hardware. We used a L14-5/38 ultrasound transducer and

acquired data with the Sonix DAQ and the MUSiiC software toolkit [15]. The third is a piezoelectric element attached to the tip of a needle. The piezoelectric element that we used was 2mm in diameter, requiring a 12 gauge needle, which we did not have available at the time. Instead, a straw was used as a temporary substitute. The main difference in using a straw instead of a needle is that it makes camera segmentation more simple. Each of these main components can be seen in figure 8.3.

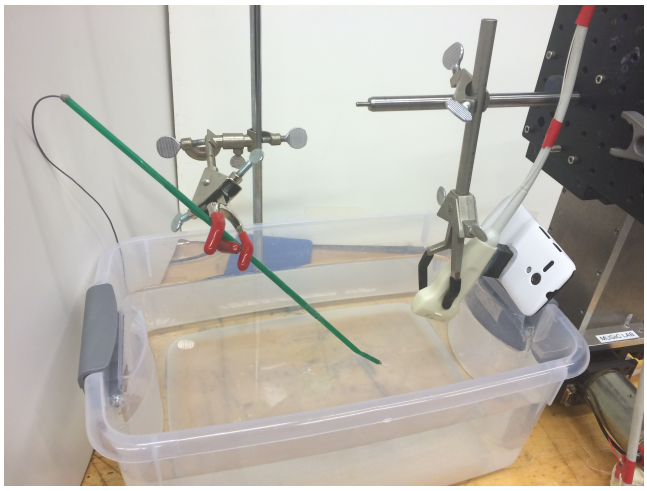


Figure 8.3: The experimental setup used to validate this sensor fusion needle tracking approach. An ultrasound transducer with an attached camera is held by a Cartesian stage as it observes a piezoelectric element.

8.4.2 Experimental Setup

Two experiments were performed to evaluate this needle tracking approach. The first involves placing the piezoelectric element at the needle tip inside of the ultrasound imaging plane. We then compare the computed position with where it appears to be in the image. The second experimental setup consists of moving the ultrasound

transducer with a Cartesian stage to a set of known locations. The piezoelectric element is then localized at each of these independent locations. This experimental setup can be seen in figure 8.3. We use the relative accuracy measure described in equation 8.4 to validate this method. N is the number of total data points (50). q is the localized piezoelectric element position and M is the known motion or distance between any particular step. Our experiments were performed in a water tank.

$$\forall i = 1 \ldots N-1 : RA = norm(q_i - q_{i+1}) - M \tag{8.4}$$

8.5 Results

The first experiment resulted in errors of 0.63mm and 0.18mm on two independent poses. In the second experiment, this approach was used to compute the piezoelectric element location for each of the Cartesian stage configurations. Figure 8.4 and figure 8.5 represent subsets of the computed points with respect to the ultrasound image plane. As one can see, the computed points look similar to the three axes of motion applied using the Cartesian stage. Quantitatively, the relative accuracy measure described in equation 8.4 had a minimum, maximum, mean, and standard deviation of 0.02mm, 2.15mm, 0.61mm, and 0.61mm.

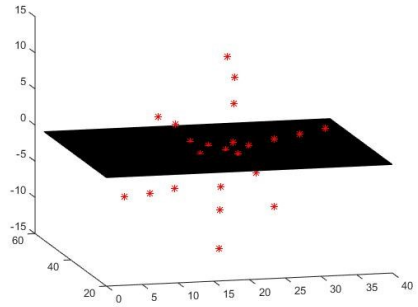


Figure 8.4: First subset of detected piezoelectric element positions with respect to the ultrasound image plane (black plane).

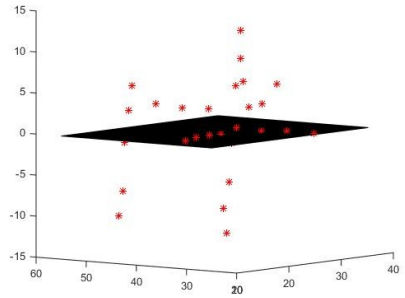


Figure 8.5: Second subset of detected piezoelectric element positions with respect to the ultrasound image plane (black plane).

8.6 Discussion

The results show that this approach is possible for detecting piezoelectric elements placed at the tip of a needle by fusing sensor information from a camera and an ultrasound transducer. There are several possible reasons for the errors. First of all, this approach relies on an ultrasound calibration process to register and relate the camera and ultrasound image coordinate systems together. Any errors in the calibration will propagate to the end tracking result. Second, we made assumptions that the needle is an ideal line and the piezoelectric element is an ideal point source at the end of this line. Any deviations of the needle, due to camera resolution, segmentation errors, or bending, will affect the result. The piezoelectric element also has some finite size, which is on the same order as the achieved errors.

Future work will look at making this tracking method real-time and also possibly explore the use of multiple piezoelectric elements to compensate for needle bending

or other uncertainties.

8.7 Conclusion

In this chapter, we demonstrated the use of sensor fusion to track a piezoelectric element with two incomplete sources of sensor information. Through experiments performed in an ideal environment, on average, sub-millimeter errors were achieved. Future work will focus on demonstrating this method in more realistic *ex vivo* scenarios and explore whether our listed assumptions hold.

Chapter 9

Virtual rigid body: a new optical tracking paradigm in image-guided interventions

9.1 Introduction

Image-guided surgery systems are frequently used during surgery to provide surgeons with informational support. For these systems to show their full capabilities and to enable more advanced applications such as volume building or automated actuation, tools and devices must be registered. An integral component to register these devices together is tracking. Electromagnetic (EM) or optical trackers are conventionally used for tracking. A sensor or marker is placed on the device, and

the tracker reports the pose of the sensor or marker. In general, conventional tracking technologies are fairly accurate and can achieve sub-millimeter tracking errors. However, both EM and optical trackers have limitations. EM sensors do not require line-of-sight, but their accuracy is compromised in the presence of metal tools. This is a fairly significant limitation while using EM trackers as ferromagnetic instruments and devices must be avoided.

On the other hand, optical trackers do not suffer from metallic distortion. The first limitation for optical trackers is that they do require line-of-sight between the optical tracker and the optical marker. Secondly, the tool size and the crowded surgical workspace limits the optical marker size. There has been some research in the design of optical markers [76]. Our work addresses the limitations of optical trackers by proposing an optical tracking paradigm using a virtual rigid body (VRB). This section describes the virtual rigid body, the algorithms to enable this technology, the experimental setup, and the results.

As shown in figure 9.1, the virtual rigid body is a 3D rigid body marker generated from a light source attached to tools. It appears as a projected pattern on some surface. Its pose can be recovered by observing the projected pattern with a stereo-camera (SC) system or a conventional optical tracker. Like conventional markers, the device projecting the VRB must be attached to the tool. While line of sight is still required between the optical tracker and the light pattern projected onto the surface, there can be built-in redundancy in the VRB so that partial occlusions of the pattern

will not prevent successful tracking. Also, the tool or device does not need to be in the stereo-cameras field of view as long as the projected pattern on the surface is visible. This can be an easy requirement to satisfy if the pattern is projected onto the region of interest guaranteed not to be occluded. The other type of occlusion that could occur is between the light projector and the surface. However, since light is being projected, some or all of the points will simply be on the surface occluding the original surface, allowing this method to work with any arbitrary surface.

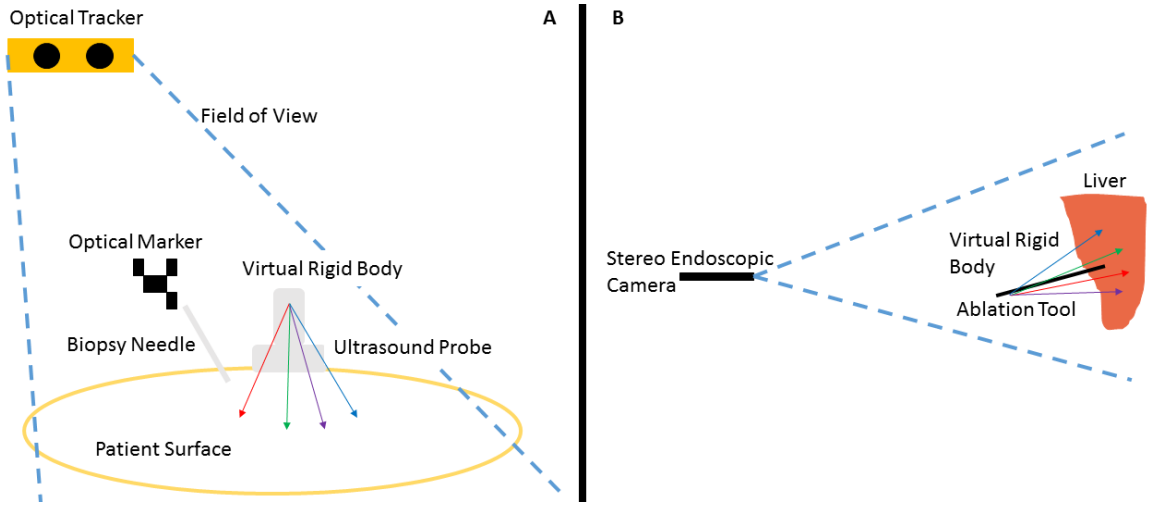


Figure 9.1: Using a virtual rigid body for tracking in A) an ultrasound guided needle biopsy and B) a minimally invasive tracked ablation clinical scenario.

The projection device that is being tracked would be attached to another medical tool or device, allowing the tool or device to be tracked after a standard calibration process. Two examples of the calibration process necessary are ultrasound (US) calibration [41] for tracked US images or pivot calibration for tracked needle tips. In conventional optical tracking, the marker attached to the tool must always be within

the optical tracker's FoV. With the VRB, the tool or device can be outside of the optical tracker's field of view (FoV) as the pattern will still be projected onto the surface regardless. Also, optical trackers have regions with varying RMS errors, so it is beneficial to minimize the size of the tracking region. The VRB allows the optical trackers to focus on a single region of interest as opposed to the region of interest and each tracked tool. Referring to figure 9.1A, the optical marker is far away from the patient surface, whereas the projected VRB is directly on the patient surface.

It is known that a larger marker will result in higher accuracy, but the surgical workspace limits the marker size. However, the size of the VRB is no longer limited by its physical size, but rather by the field of view of the optical tracker. This is one advantage of using the VRB, as it is feasible to create a large marker in a limited surgical workspace. It can be seen in figure 9.1A and figure 9.1B that the projected VRB is much larger than the optical marker. Theoretically, this type of optical tracking could allow for a smaller form factor than conventional tracking methods. Also, the VRB can be customized to project any pattern. For example, if the tracked tool is likely to be far away from the region of interest, then it may be preferable to have a VRB with a more compact pattern such that most of the pattern is within the optical tracker's FoV.

To convey the use of the VRB with a more concrete example, we will describe the clinical scenario shown in figure 9.1B. During a minimally invasive guided liver ablation, the surgeon will have access to a stereo-endoscopic camera and an ablation

device. The device that generates the VRB will be rigidly attached to the ablation tool, projecting some pattern onto the liver surface. It is assumed that the rigid body transformation between the VRB's coordinate frame and the ablation tool tip is found pre-operatively with a calibration process. The stereo-endoscopic images are processed to track the VRB and consequently the ablation tool. These stereo-endoscopic images could also be used to digitize the liver surface and register it to some pre-operative ablation plan. Visual cues can now be displayed to the surgeon to help them position and orient the ablation tool when executing their pre-operative plan.

Our method shares some similarities with structured light systems [77, 78]. Both methods rely on light projection devices to project some known pattern onto an arbitrary surface. However, the goals are different as structured light systems aim to find a 3D representation of the surface that it is projected on, whereas our system seeks to find the pose of the light projection device. The technology developed for structured light systems such as automatic detection and correspondence recovery [78] could also be applied to this method.

There has also been work in the virtual reality field on the design of six degree of freedom tracking devices. Systems such as Kinectrack [79] and Sceptre [80] track projected patterns of light on planar surfaces to deduce the six degree of freedom pose of the emitter. These works assume that the surfaces are either planar or piecewise-planar. This is a valid assumption in virtual reality environments as there are either

a large number of projected points or the surface is a planar screen. However, in image-guided interventions, there are rarely planar surfaces in the human body. As such, it is necessary for a method to track the projection device based on a set of projected points on an arbitrary surface. To the best of our knowledge, this is also the first application of projected light as a tool tracking method in medical applications.

9.2 Contributions

The main contributions of this chapter are the following:

- Development of an optical tracking method and apparatus based on a non-physical fiducial made of light
- Experimentation to compare VRB optical tracking with conventional optical tracking
- Demonstration of targeting errors and tracking errors comparable to conventional optical tracking

9.3 Acknowledgement

Dr. Nishikant Deshmukh and David S. Lee were also involved in this work. I developed the initial prototype with Dr. Nishikant Deshmukh during a class project. I led the algorithmic design and the experimental validation. I then further developed the theoretical background before mentoring David S. Lee when we developed the

second prototype of this project. I led the experimental design and validation.

9.4 Technical Approach

9.4.1 Calibrating the virtual rigid body

As was previously mentioned and shown more concretely in figure 9.2, the virtual rigid body is a 3D rigid body marker and is generated from a light source. Some light source, attached to the tool, such as laser lights, LED lights or a small pico-projector are used to project a pattern onto a surface. While figure 9.2 shows the VRB projected onto a planar surface, we will describe in details the algorithms that the surface can be arbitrary. Figure 9.2 also shows two different kinds of patterns, checkerboard and points, being projected onto the surface. They can be interpreted as the same type of pattern for the purposes of our algorithm because each vertex or cross point in the checkerboard can be interpreted as a point. In this work, we used a pico-projector as our light source, projecting a checkerboard pattern onto the surface. Projecting a checkerboard pattern also allows for surface reconstruction.

While the projection pattern is theoretically known since we control the projected image, we decided to explicitly calibrate the projection pattern using a stereo-camera system. First, we fixed the projector and the stereo-camera system. Then, we take stereo-camera images as the VRB is being projected onto multiple surfaces. Each of the checkerboard cross-points represented as three-dimensional coordinates in the

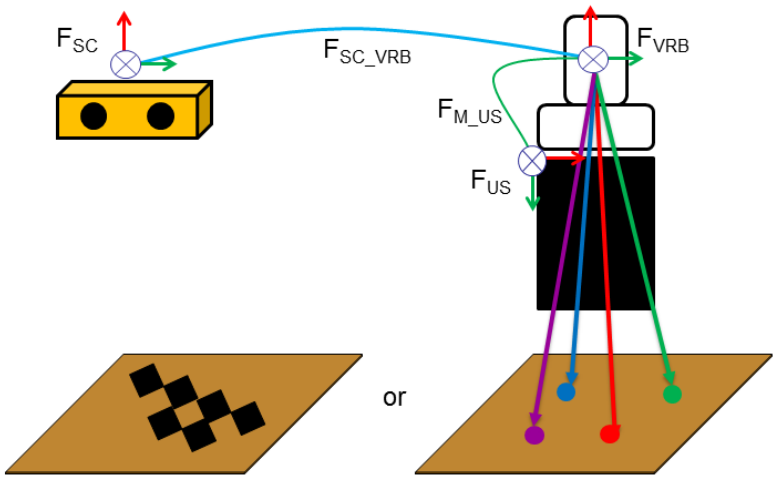


Figure 9.2: A conceptual figure demonstrating the use of a VRB for tracking an ultrasound transducer. Two possible VRB projection patterns are shown.

stereo-cameras coordinate frame are corresponded with their respective checkerboard cross points in subsequent stereo-camera images. We then applied a best-fit line to each of the checkerboard cross-points, resulting in the projection vectors in the stereo-cameras coordinate frame. The final step is to redefine these projection vectors in the VRBs local coordinate frame. At minimum, stereo-camera images from two different surfaces are necessary to determine the projection vectors.

9.4.2 Tracking with the virtual rigid body

There are three main components to enable tracking with the VRB. The first component is segmenting the checkerboard cross points from the pair of stereo-camera images. Since we are using a MicronTracker (Claron Technology) as our stereo-camera system, it has native support for automatically segmenting checkerboard cross points.

Currently, this automatic segmentation is limited to planar and piece-wise planar surfaces. However, there are many methods in the structured light field that can robustly detect grid patterns on arbitrary surfaces.

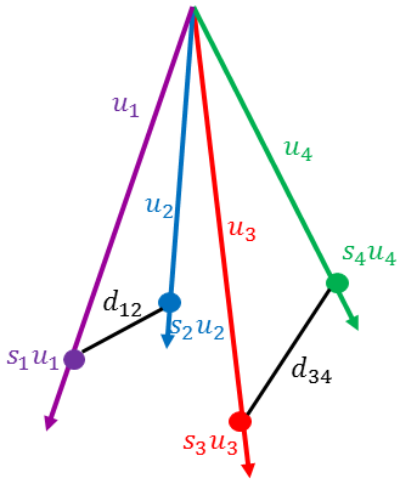


Figure 9.3: Notation and parametrization of the projection pyramid used for virtual rigid body tracking. Its basic embodiment consists of four projection rays.

The second component is matching the checkerboard cross points observed by the stereo-camera to the projection pyramid. We demonstrate this by using the pyramid shown in figure 9.3. The four lines represent the projection vectors, fully described by u_i, connecting the projectors optical center and each checkerboard cross point. The number of lines shown is simply an example and there is no theoretical limit to the number of projection vectors. The apex of this pyramid is defined as the origin of the projection pyramids coordinate system. These lines can be parameterized such that any point on the line is defined by a scale factor, s. Thus, given an observation of the surface and projection vector intersection points p, in the stereo-cameras coordinate system, we wish to recover the corresponding points on the projection pyramid. Liu et

al. [81] previously used a similar approach to recover the pose of a plane intersecting such a pyramid.

$$\forall i, j \; and \; i \neq j \quad d_{ij}^2 = |(p_i - p_j)|^2 \tag{9.1}$$

$$\forall i, j \; and \; i \neq j \quad d_{ij}^2 = |(s_i u_i - s_j u_j)|^2 \tag{9.2}$$

$$\forall i, j \; and \; i \neq j \quad d_{ij}^2 = s_i^2 + s_j^2 - 2 s_i s_j cos(\theta_{ij}) \tag{9.3}$$

To show that this approach is feasible, we must first define a set of d_{ij} using equation 9.1, which represents the distances between the checkerboard cross points in the stereo-cameras coordinate system. We must then solve the non-linear optimization problem described in equation 9.2 using methods such as trust region reflective [82] or Levenberg-Marquardt [83, 84], where we wish to find the set of s. Assuming that the set of projection unit vectors, u, are unique, it is clear that a solution s can be obtained from equation 9.2. This is a valid assumption as we have full control over the choice of each u_i. However there is still an uncertainty whether the solution s is unique given a set of d_{ij}.

Equation 9.2 can be rewritten as equation 9.3 where θ_{ij} represents the angle between u_i and u_j. We will show that the solution s is unique when there are three orthogonal projection vectors, u. When this is the case, the $cos(\theta_{ij})$ term in equation 9.3 becomes zero. The set of equations then devolve into a simple quadratic form where one can isolate s_i and solve the system of equations without using any optimization algorithms. Due to the equations being quadratic, each s_i will have a

positive and negative solution. In our parameterization, it is impossible for s_i to be negative as the light is only projected in one direction. When the projection vectors are not orthogonal, then it becomes possible for s_i to have multiple positive solutions, leading to a non-unique solution. In these cases, it is necessary to add another projection vector. Thus, the minimum number of projection vectors with arbitrary directions to enable optical tracking is four.

The third component is to find the rigid body transformation between the checkerboard cross points defined in the stereo-camera and the projection pyramid coordinate system. There are a number of different algorithms that can be used for this purpose such as Aruns method [85] or Horns method [86], which we used. This results in the pose of the projection pyramid or VRB in the stereo-cameras coordinate system, allowing us to track the VRB.

9.5 Methods

9.5.1 Apparatus

There have been two iterations of the apparatus used to generate a virtual rigid body. The first generation shown in figure 9.4 consisted of laser pointers physically arranged such that the beams of light formed a projection pyramid. The laser pointers are fixed by two acrylic plates and are well below the safety limits for applying laser energy to human tissue. The laser pointers are also chosen to emit different

Figure 9.4: Notation and parametrization of the projection pyramid used for virtual rigid body tracking. Its basic embodiment consists of four projection rays.

wavelengths of light. In this case, the laser pointers projected red, blue, and green light respectively. This allows the correspondence between the points seen by the SC and the projection vectors of the projection pyramid to be recovered.

There were several limitations to this first generation virtual rigid body. First, the projection vectors could not be reconfigured without a new set of acrylic plates to hold the laser pointers in a different orientation. It would be advantageous if the projection vectors could be customizable as different configurations or number of projection vectors may be beneficial in various situations. Second, each spot related to each projection vector could not be changed to a different pattern, such as a cross. The ability to do so would aid the development of automatic segmentation algorithms as specific patterns can be easier to segment than the point projected by

a laser pointer. Also, since it was made with three projection vectors, we had to re-project each of the viable solutions into the stereo-camera image to determine the correct solution.

To address the limitations listed above, the second generation used a Microvision Showwx projector (Microvision, Inc., Washington) in place of laser pointers. Since a projector simply projects an image, the user now has complete control of the projected pattern. Different projection pyramids and spot patterns are easily configurable. Additionally, custom acrylic plates do not need to be manufactured. A projector can be used as an optical tracker using these methods out of the box.

In addition to the projection device, the concept of optical tracking with a virtual rigid body also requires a SC system. In our experimental setup, a MicronTracker Sx60 (Claron Technology, Toronto, Ontario, Canada) is used as the SC system. In this setup, we can leverage the marker detection algorithms present within the MicronTracker software development kit. A typical SC marker is also attached to the projector so that a comparison between conventional optical tracking virtual rigid body optical tracking can be made. The field of view of these cameras is approximately 40°.

9.5.2 Experimental Setup

We designed an experiment to test the feasibility of optical tracking with a VRB. This experiment compares the motion tracking accuracy of our method and conven-

tional optical tracking. In this case, we separately used a linear stage and a Universal Robots UR5 robotic arm to generate the ground truth. We use the linear stage because it has better accuracy than the robotic arm, while the robot arm gives us more precise complex motions like rotations. We make known rotational and translational motions and compare the VRB with the known motions. We repeat this with the robot arm holding a conventional optical marker for the MicronTracker instead.

In the first experiment using a MicronTracker and a linear stage, acquire the set of SC images and the known translation information at a number of different poses. After computing the projector pose from the set of SC images, we can check the difference between the computed and known translation for every pair of poses. It should be noted that in this first experiment, the conventional optical marker consisted of only a single checkboard pattern and can therefore only recover position.

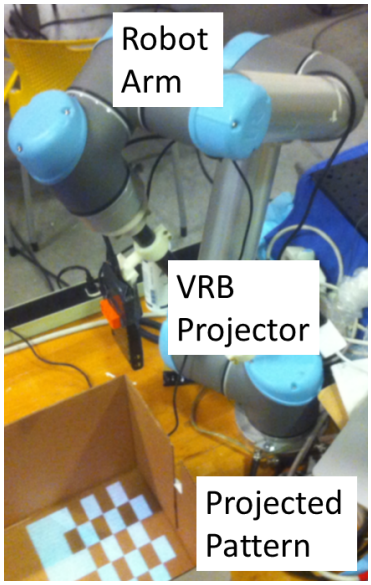


Figure 9.5: Experimental setup of the virtual rigid body projector held by a robotic arm, projecting a checkerboard pattern onto a surface.

As shown in figure 9.5, the second experiment replaces the linear stage with a Universal Robots robotic arm to allow for rotations. The robotic arm is rated to have $100\mu m$ precision. Precision is sufficient for our purposes because we are dealing with relative accuracy metrics. The metric used for the second experiment is derived from the magnitude of the motion. Each motion can be represented as a transformation matrix composed of a rotation matrix and a translation vector, $F = [R, t]$. We can compare d when the motion is translation only and θ when the motion is rotation only, where these two values are computed using equation 9.4 and 9.5 respectively. In equation 9.5, $e^{\theta N}$ denotes the matrix exponential and N is the skew symmetric matrix generated from the vector representing the axis of rotation. These two values respectively represent the magnitudes of a translation or rotation motion.

$$d = norm(t) \tag{9.4}$$

$$R = e^{\theta N} \tag{9.5}$$

In terms of data collection, both of these experiments used motions generated about the robots main principal axes. This includes translation motions parallel and normal to the robot base and rotation motions about the robot bases three main axes. Motions about arbitrary axes were not collected in these experiments. However, since we are dealing with relative motions generated from any two poses, there are many resulting relative motions about arbitrary axes used in the experimental analysis.

Type of Motion	VRB	Conventional
Translational (mm)	0.18±0.10	0.03±0.02
Rotational (°)	0.50±0.31	2.68±2.20

Table 9.1: Error metric comparison of VRB and conventional optical marker in two motion scenarios

9.6 Results

The first experiment was done using a MicronTracker and a linear stage, resulting in targeting errors of $0.38mm \pm 0.28mm$ with our method and $0.23mm \pm 0.22mm$ with conventional optical markers. The results for the second experiment can be seen in table 9.1.

9.7 Discussion

There are several observations to be made from the results. First of all, the errors of our new optical tracking method are comparable to conventional optical tracking. We can see that, in the first experiment with the linear stage, the translation errors using the virtual rigid body are higher than when using conventional optical tracking. This can be seen even more evidently in the second experiment. One likely reason is that the MicronTracker is optimized to segment the attached markers. At the same time, we also see that the rotational errors are much better when using the VRB than when using the conventional optical marker. One possible reason for

this is that the VRB has better rotational sensitivity. Any rotation experienced at the tool will have its effects magnified on the projected pattern. More experiments must be conducted to evaluate when the size of the virtual fiducial overcomes its inherent disadvantage in segmentation. One result that seems slightly odd is that the translational error using a conventional marker is even lower than the rated robot arm precision. As we previously mentioned, the collected motions were about the robots principal axes. This means that the change in gravity and the associated gravity compensation mechanisms are minimal. Thus, the robot arms precision may have been better than the rated value in the motions that we selected to use.

There are several advancements that must be made before this technology can be brought from the bench-top to a practical setting. A primary requirement will be miniaturizing the projection system. The projector used is evidently inadequate for a real application as it requires wired power and a processor. However, it serves as an ideal prototype as it gives us flexibility in the projection pattern. Also, the projector that we are currently using is much too large and impossible to attach to the tool or device that we wish to track. Some possibilities would be to use a fiber delivery or laser diode system. Another possibility is to use a light emitting diode with a mask to generate a pattern. Another advancement may be an alternate light source that emits infrared light. This combined with a SC system that receives only infrared light may significantly aid in automatic segmentation.

Optimized automatic segmentation is another barrier to the eventual deployment

of the system. The standard automatic segmentation present in the MicronTracker fails when the projection of the checkerboard grid becomes irregular due to the surface. There are some methods [78] in the structured light field that can be used to robustly detect these grids and also establish correspondence. This type of pattern would also allow us to recover the 3D representation of the surface, which could be one possible method to reference the patient.

One other challenge for optical tracking methods in the surgical environment is the lighting. Lighting is an important consideration if one wishes to use the intensity and wavelength of the projected light for segmentation and correspondence. One can either use wavelengths that do not interfere with the surgical lighting or use structured light approaches that depend less on wavelength. One related challenge in establishing correspondence is when multiple tracked tools are present. One could use coding strategies where the patterns are independently or selectively turned on or off. Wavelengths that are distinguishable from each other could also be used.

In the future, VRB optical tracking can be used in many applications that require this technology. Examples may include online tracked ultrasound elastography (O-TRUE) [87] and integration of tracked elastography with robotic surgery [88,89]. Each of these applications can use the VRB to enhance tracking for detecting in-plane US slices for higher quality elastography images.

9.8 Conclusion

In this work, we showed the feasibility of optical tracking using a virtual rigid body. Given the same optical tracking equipment and equivalent segmentation methods, this new optical tracking paradigm performs comparably with conventional optical tracking in terms of targeting error. Future work will include extending automatic segmentation for arbitrary surfaces and miniaturizing the prototype for practical use.

Chapter 10

Photoacoustic-based catheter tracking simulation, phantom, and *in vivo* studies

10.1 Introduction

Laparoscopic procedures have become common practice as the demands for shorter hospital stays and faster patient recovery times have increased [90]. This has created a need for the tracking of tools and devices, such as catheters, within the patient body. Catheters are used in many applications such as cardiac catheterization. In this procedure, a catheter is inserted into the body and is guided manually towards the target through the patients venous system. One complication that may occur is

that the catheter is inserted into an incorrect vein due to a lack of real-time guidance.

Catheters are commonly tracked with either some kind of imaging modality such as fluoroscopy or ultrasound (US) or some kind of tracking system such as electromagnetic (EM) trackers. X-ray fluoroscopy-based systems [91] are the most common, but they have a very obvious disadvantage. X-ray fluoroscopy delivers radiation dose to the patient for a non-therapeutic purpose. Due to its harmful nature, its use for catheter tracking is also diminished because it will generally not be used in a real-time and continuous manner. The general use-case scenario is to take intermittent x-ray fluoroscopy images as the catheter is manually inserted. On the other hand, EM-based catheter tracking systems [92] can provide real-time tracking. In these systems, an EM sensor will be placed on the catheter, while a field generator is placed external to the patient. As long as the EM sensor is within the field generators tracking volume, real-time catheter tracking can be achieved. However, EM tracking places certain restrictions on the types of tools that can be used during the procedure, as ferromagnetic materials affect the tracking accuracy significantly.

In this section, we present a novel catheter tracking technology that makes use of photoacoustic (PA) imaging and more specifically, the photoacoustic effect. Photoacoustic imaging is an emerging imaging modality that captures information about the optical properties of the target material or medium. Its capabilities are based on the photoacoustic effect, which can be described as a conversion from light to acoustic pressure. This acoustic pressure can then be received by a conventional ul-

trasound transducer or a piezoelectric (PZT) element. Thus far, photoacoustic imaging has been mainly used as an imaging modality for interventions such as prostate brachytherapy [65, 93]. Its use as a tracking technology has been fairly limited. We previously demonstrated methods in part II for tracking ultrasound transducers using photoacoustic markers, points that can be simultaneously observed optically and acoustically [30]. A catheter tracking method using photoacoustics would increase the attractiveness of photoacoustic systems by providing them with the means to be an imaging and general tool tracking solution.

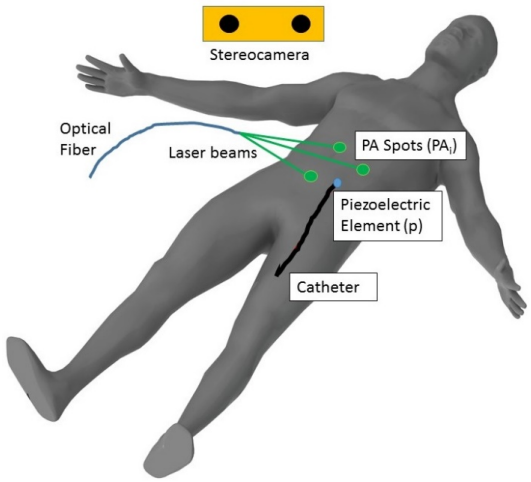


Figure 10.1: Photoacoustic catheter tracking concept figure. Multiple photoacoustic spots are projected onto the surface of the patient body, generating an acoustic signal due to the photoacoustic effect. A stereocamera and piezoelectric element can simultaneously capture data related to these spots.

Figure 10.1 is the concept figure for our photoacoustic catheter tracking work. We leverage the use of photoacoustic markers, denoted as PA_i. A stereo camera (SC) is used to observe these markers optically. These markers are non-physical entities generated on the surface of a medium. A piezoelectric element is also attached to

the catheter tip, denoted as p, and acts as an acoustic sensor. The signal received by the PZT element for each marker will correspond to a distance measure between the marker and the PZT element itself. As long as there are more than three well-structured photoacoustic markers, the catheter tips location with respect to the stereo camera system can be computed. This method addresses the limitations described for x-ray fluoroscopy-based and electromagnetic-based catheter tracking systems, but it does place a constraint that the stereo camera system must have line of sight to the surface.

10.2 Contributions

The main contributions of this chapter are the following:

- Development of a photoacoustic-based method for catheter tracking
- Experimentation in phantom, simulation, and *in vivo* scenarios
- Demonstration of error metrics that may be suitable for certain surgical procedures

10.3 Acknowledgement

I wish to thank Yuttana Itsarachaiyot who was involved in the early parts of this project and some of the initial experiments, Younsu Kim for his help with the PZT system, and Dr. Clifford Weiss for his help with the *in vivo* experiment. I led all

components of this project.

10.4 Technical Approach

This photoacoustic catheter tracking method has four main components: data collection, stereo camera segmentation, piezoelectric signal segmentation, piezoelectric element localization.

10.4.1 Data Collection

In our method, there are two sources of data that must be synchronized and collected. The first source is from the stereo camera system. Naturally, the images acquired must by synchronized with the light source. In this case, we use the external hardware trigger from the laser system to trigger image capture. This external hardware trigger is also used to trigger PZT signal collection. For the time being, we collect data from a single photoacoustic marker at a time. This keeps the PZT signals independent and not accumulated from multiple photoacoustic markers.

10.4.2 Stereo-camera Segmentation

Stereo camera segmentation is the process that takes the images acquired of a single photoacoustic marker and recovering the three-dimensional position of this marker within the stereo cameras coordinate system. This requires the point to first

be segmented in each camera image. To accomplish this, we use intensity-based and shape-based filters. To further aid this process and avoid false positives, we remove the background in the camera image by subtracting an image that does not contain the signal. Further details of this segmentation process can be found in chapter 4. After this is completed for each of the camera images, the point can be triangulated into a single three-dimensional point, denoted as PA_i.

10.4.3 Piezoelectric Signal Segmentation

A PZT element receives this time-synchronized acoustic signal. Since the collection of this acoustic signal is synchronized with the laser, the location of the signal in this waveform represents the time between the laser firing event and the acoustic signal reaching the PZT element, denoted as ToF_i. A band-pass filter is used to remove noise in cases where the acoustic signal may be weak. Following that, we use a peak detector to determine the aforementioned time between the laser firing event and the acoustic signal reaching the PZT. A sample of a collected photoacoustic signal can be shown in figure 10.2.

10.4.4 Piezoelectric Element Localization

The previous components obtained a three-dimensional position for each photoacoustic marker as well as a measure of time between the PZT element and each of

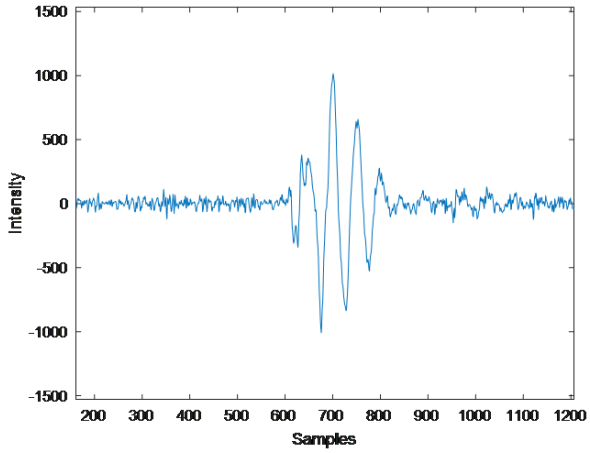


Figure 10.2: Sample acoustic signal received by the piezoelectric element

these positions. Assuming that the medium has a homogeneous speed of sound, SoS, each of these measures of time can be converted to distance. This forms a trilateration or multilateration problem. One can then write a distance equation as shown in equation 10.1. With multiple PA spots, this becomes a quadratic system of equations and can be solved when there are at least 3 non-collinear and well-structured PA spots. In the case of three PA spots, there will be two solutions where the piezoelectric element is either above or below the surface, with the latter being the desired solution. This ambiguity can be easily solved as the point below the surface will always be farther away from the stereo camera system than the one above the surface. A system with more than three PA spots will require optimization.

$$\forall i = 1 \ldots n : \underset{p \in R^3}{\arg\min}(||PA_i - p||_2 - ToF_i * SoS)^2 \tag{10.1}$$

Piezoelectric element localization can also be shown graphically in figure 10.3. Each of the photoacoustic markers are located somewhere in three-dimensional space. Since the PZT element must lie at a known distance from each of the photoacoustic markers, one can draw a sphere with the distance being its radius, centered about each photoacoustic marker. With three markers, there will be two solutions where each of the three spheres intersect.

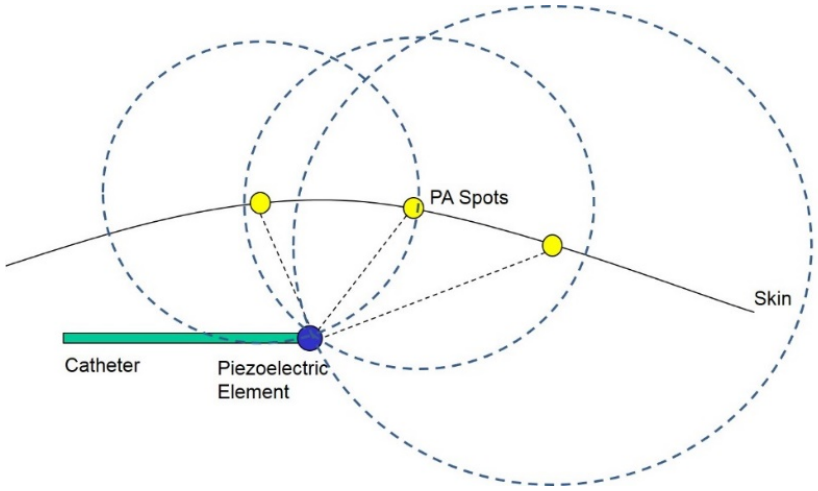


Figure 10.3: Graphical description of piezoelectric element localization

The localization methods presented in this section do not attempt to accommodate for errors. We rely on least-squares minimization to reduce the impact of errors. One possible next step is to integrate errors into the localization methods. For example, if we consider the graphical representation shown in figure 10.3, each of these spheres would have a shell with some thickness correlated with the magnitude of the errors. Thus, there could be a set of possible solutions present in the overlapping volume of

these spheres. One could envision integrating probabilistic methods to find the most likely solution given some initial knowledge of the error model.

10.5 Methods

10.5.1 Phantom

The experiment was performed with a plastisol phantom. The laser system used to generate the photoacoustic markers is Q-switched neodymium-doped yttrium aluminum garnet (Nd:YAG), Brilliant (Quantel Laser, France) laser, operating at 532 nm and approximately 1 mJ/cm^2. This energy density is well below the safety limits [7]. A stereo camera system consisting of two CMLN-13S2C cameras (Point Grey Research, Canada) is used to acquire the camera images. We used a PZT element as shown in figure 10.4. It is made of a customized PZT5H tube with an outer diameter of 2.08mm, an inner diameter of 1.47mm, and a length of 2mm. The signal from the PZT element is collected using one of the channels of a data acquisition system, the Sonix DAQ (Ultrasonix, Canada) and associated software [8].

The experiment was designed to test the repeatability of this PZT element localization method. As shown in figure 10.5, the entire setup is placed at 4 known locations. At each location, 20 photoacoustic markers were collected. Ideally, any three of these photoacoustic markers can be used to localize the same PZT element position. Naturally, there must be some additional constraints on which of these pho-

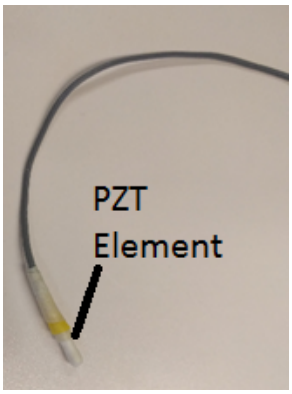


Figure 10.4: Experimental setup and apparatus

toacoustic markers are selected, such as non-collinearity and total area. With this data, we were able to analyze a precision metric as well as a relative accuracy metric.

Another experiment was performed on the same apparatus at a single location. A much denser sampling of points was acquired, totaling 65 photoacoustic markers across 4 linear motions. The data from this experiment can be seen in figure 10.6. The red spots in the left image correspond to the segmentation result. Some rejection was necessary for weak observed signals. From figure 10.6, one can see that the general trend of the observed acoustic signal agrees with our expectation based on the PA spot motion. This experiment was designed to analyze the effects that the number or dimensions of photoacoustic markers has on the proposed method. To validate this, we use a precision metric, a leave-out metric, and an estimated accuracy metric. Each of the metrics for our experiments will be mathematically defined in the following section.

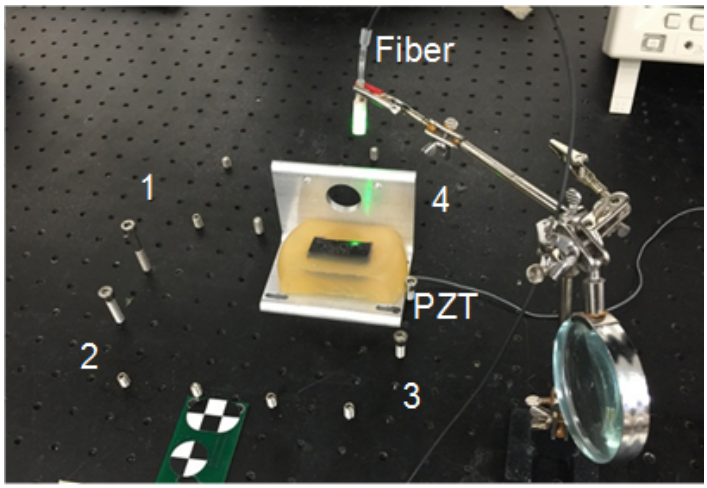


Figure 10.5: Experimental setup and apparatus

10.5.2 Simulation

Simulations were conducted to examine the effects of four parameters: source radius, sensor radius, sensor length, and speed of sound. Simulations were done using the k-wave toolbox [67]. Four sources and 1 sensor were placed in the arrangement shown in figure 10.7. An accuracy measure comparing the computed location with the known location is used for validation.

10.5.3 *In Vivo*

An *in vivo* porcine was performed to evaluate this system in a realistic environment. To enable real-time data collection, the apparatus used was slightly different than in the phantom case. There were two main differences: a galvonometer to actively control where the PA spot was and an oscilloscope for piezoelectric sensor data

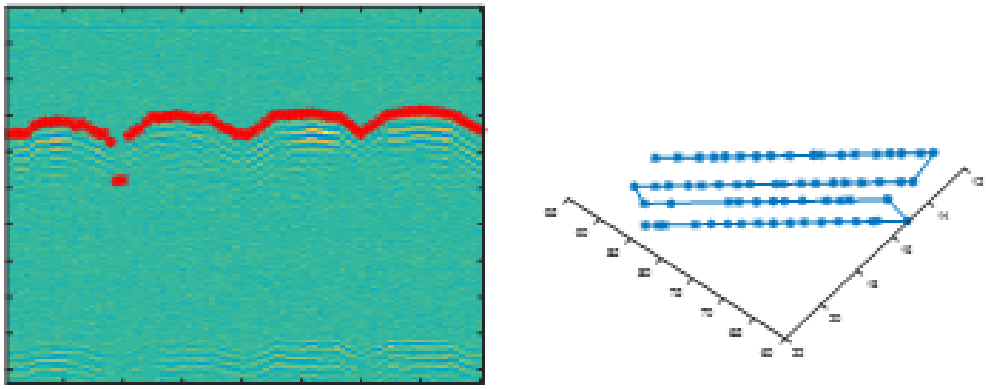


Figure 10.6: Left) Observed acoustic signal by the piezoelectric element during PA spot motion annotated with segmentation result. Right) PA spots in the stereocamera space across 4 linear motions

capture. Together, these changes gave us much higher data collection rates than the original phantom experiment setup. This setup can be seen in figure 10.8.

Another key is the use of a photoacoustic active layer, which we put on top of the pig's skin surface. This gives us control over the generated photoacoustic signal. In addition, we place fiducials on this layer, letting us register our camera tracking system with computed tomography (CT) data for validation purposes. These fiducials also enable surface tracking with the cameras for possible motion compensation due to respiratory activity.

The procedure was as follows. Our piezoelectric element tool was placed inside a 9 french catheter. A surgeon then navigated the catheter to a renal vein under the guidance of ultrasound and fluoroscopy. Afterwards, we use our method to compute the location of the piezoelectric element and compare it with the element segmented from a CT volume.

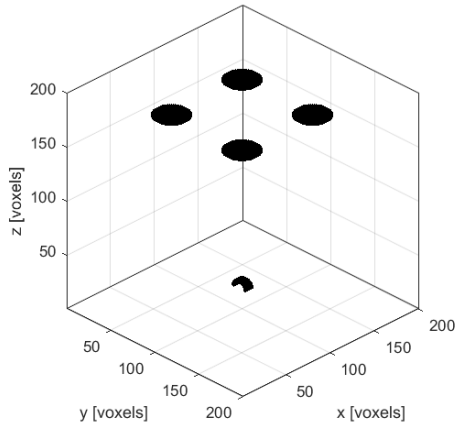


Figure 10.7: Arrangement of sources and sensor in simulation environment

10.6 Results

10.6.1 Phantom

Several metrics were used to validate our experiments. These include: precision, relative accuracy, leave-out accuracy, and estimated accuracy. For simplicity, we will define several functions used in the metrics. Let PA_i and ToF_i represent a subset of the collected data. Then, our proposed method is a function, that uses these two subsets to compute the position of the PZT element, p. This can be written as $p = multilateration(PA_i, ToF_i)$. Each subset of data used will result in a different p being computed, so we can index p as p_j where j corresponds to a single trial. Over a number of trials, a mean, $\bar{p}$, can be computed. Reconstruction precision is defined as shown in equation 10.1 and equation 10.2. It represents the mean and standard

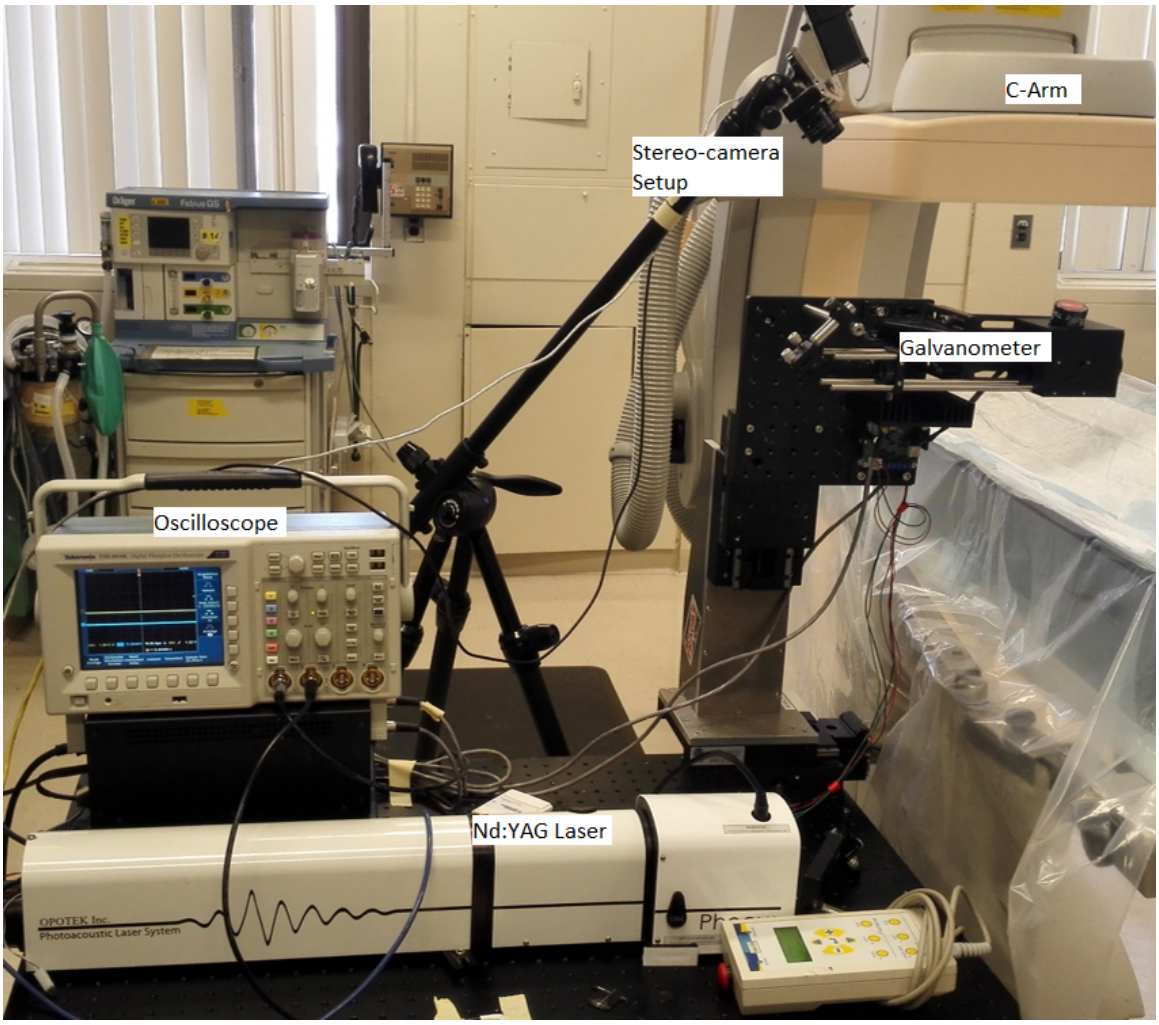


Figure 10.8: *In vivo* experimental setup including a laser, a galvonometer, and an oscilloscope

deviation of the computed points to the mean of all computed points.

$$\forall j = 1 \ldots n : distToMean(j) = norm(p_j - \bar{p}) \tag{10.2}$$

$$RP = mean(distToMean) \pm std(distToMean) \tag{10.3}$$

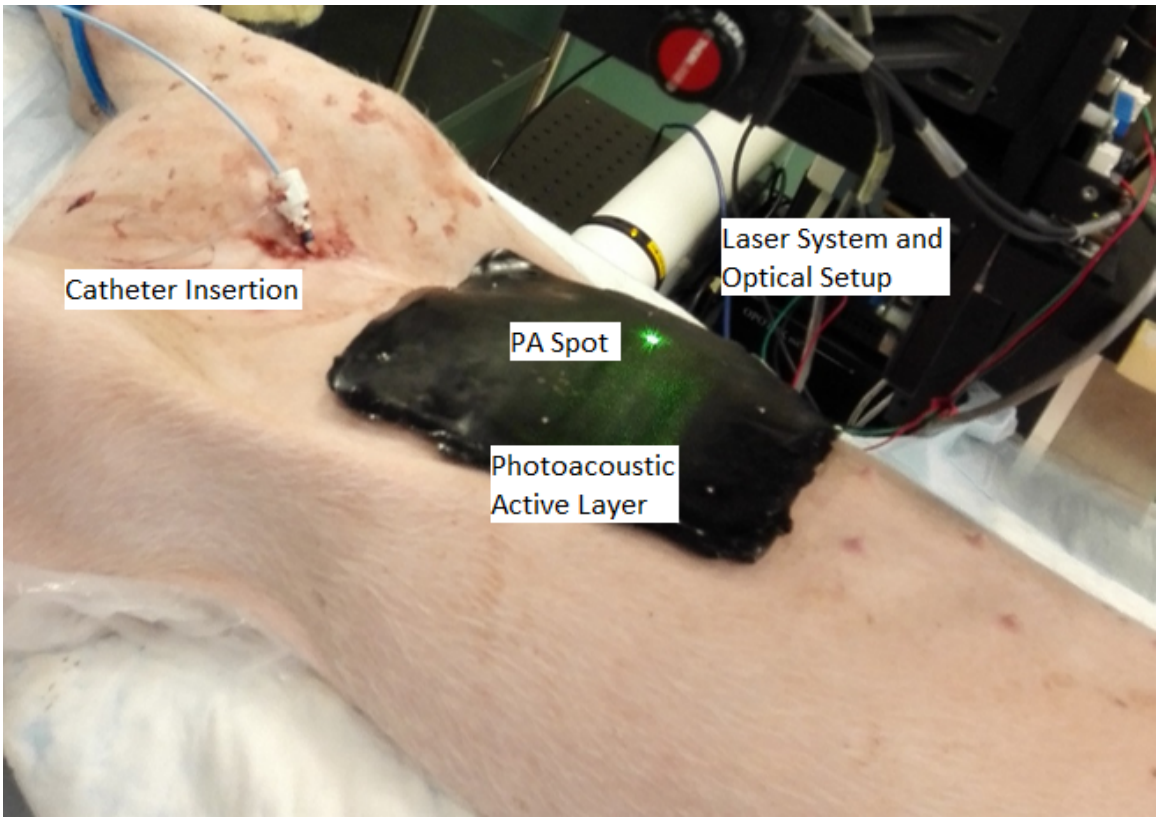


Figure 10.9: Photoacoustic active layer placed on top of pig during *in vivo* experiment

Since the PZT element is moved with known motions, one can define a relative accuracy metric. In this case, $\bar{p}$, will require an index, k, for each location. Relative accuracy is defined as shown in equation 10.4 and equation 10.5. It represents the difference between the distance of the mean computed point at 2 locations versus the measured distance. The reconstruction precision and relative accuracy for our first experiment can be seen in table 10.1 and table 10.2.

$$calcDist(k_1, k_2) = norm(\bar{p_{k1}} - \bar{p_{k2}}) \tag{10.4}$$

Location	Reconstruction Precision (mm)
Left Bottom (LB)	1.60±1.13
Left Upper (LU)	1.46±0.94
Right Bottom (RB)	2.14±1.46
Right Upper (RU)	3.12±2.43

Table 10.1: Experiment 1 reconstruction precision

Locations	Calculated Distance (mm)	True Distance (mm)	Error (%)
LB & LU	69.93	71.12	1.71
RB & RU	68.43	71.12	3.93
LB & RB	96.66	93.47	3.30
LU & RU	96.30	93.47	2.94

Table 10.2: Experiment 1 relative accuracy

$$RA(k_1, k_2) = \frac{abs(calcDist(k_1, k_2) - measDist(k_1, k_2))}{measDist(k_1, k_2)} \tag{10.5}$$

Leave-out accuracy uses the observed data that was not used to compute the PZT location as a means for validation. As we mentioned previously, each computed point uses a subset of the collected data. The remaining data can then serve as test points for validation. This metric can be seen in equation 10.6 and equation 10.7. It takes the computed point for each subset of data and finds its distance to each PA_i not within the subset. This can then be compared directly to the acoustically measured

distance, $ToF_i * SoS$.

$$\forall i, j : diffDistToTestPoints(i, j) = abs(||PA_i - p_j||_2 - ToF_i * SoS) \tag{10.6}$$

$$LA = mean(diffDistToTestPoints) \pm std(diffDistToTestPoints) \tag{10.7}$$

Estimated accuracy is a measure to compare the computed PZT location with the estimated ground truth point. Ideally, one would have a ground truth point to directly compare to the computed PZT location. One possible way of obtaining an estimated ground truth point is through the use of an independent imaging modality such as CT. This is what we use for our *in vivo* experiments. However, we use an alternative for our phantom experiments due to difficulties with making the bench setup compatible with the CT imaging system. We use an estimated one instead based on a consensus from all of the data. This consensus, q, is computed using our multilateration method with all of the data as the input. Estimated accuracy is then the difference between each computed PZT location and the consensus, q, and is shown in equation 10.8 and equation 10.9.

$$\forall j : distToGroundTruth(j) = norm(p_j - q) \tag{10.8}$$

$$EA = mean(distToGroundTruth) \pm std(distToGroundTruth) \tag{10.9}$$

In addition, we describe a partition and span concept as 2 parameters to investigate how the size or number of PA spots affects the aforementioned metrics. Both of these concepts are based on dividing the available points into different regions, then selecting a point randomly from each region. The partition parameter refers to the number of divisions each of the motion paths are split up equally into. For example, if the partition parameter is 2, then each of the 4 motion paths are divided into 2 into a total of 8 regions. A single point is randomly chosen from each region, resulting in a total of 8 points being used for localization. Partitioning will allow us to see the effects that the number of PA spots have on the aforementioned metrics.

Span is in essence a subset of the set of partitioned regions. We fix the number of partitions to a constant and pick a subset of them. The selection of this subset is chosen by the span parameter. An example can be seen in figure 10.10. Each motion path is split up into 8 partitions, with them being numbered incrementally from the center of each motion path. The numbered regions matching the span parameter will be the ones where a single point is randomly selected from each of them. We can see that the span parameter is analyzing the effects of the size of the PA spots, while keeping the number of PA spots fixed. The results for the second experiment analyzing the reconstruction precision, the leave-out accuracy, and the estimated accuracy under various partition and span parameter values is found in table 10.3 and table 10.4.

4	3	2	1	1	2	3	4
4	3	2	1	1	2	3	4
4	3	2	1	1	2	3	4
4	3	2	1	1	2	3	4

Figure 10.10: Sample description of span

Partition	Reconstruction Precision (mm)	Leave-out Accuracy (mm)	Estimated Accuracy (mm)
2	1.07±0.51	2.38±1.33	1.25±0.74
3	0.82±0.40	2.33±1.32	0.95±0.54
4	0.70±0.35	2.32±1.32	0.81±0.48
5	0.58±0.29	2.28±1.31	0.67±0.40
6	0.45±0.22	2.21±1.30	0.52±0.41

Table 10.3: Experiment 2 Partition Analysis using Reconstruction Precision, Leave-out Accuracy, and Estimated Accuracy

10.6.2 Simulation

The first parameter varied was the source radius. The evolution of this parameter can be seen in figure 10.11. The results can be seen in table 10.5.

The second parameter varied was the sensor radius. The evolution of this parameter can be seen in figure 10.12. The results can be seen in table 10.6.

The third parameter varied was the sensor radius. The evolution of this parameter can be seen in figure 10.13. The results can be seen in table 10.7.

The fourth parameter varied was the assumed speed of sound. The true speed of

Span	Reconstruction Precision (mm)	Leave-out Accuracy (mm)	Estimated Accuracy (mm)
1	0.40±0.15	2.55±1.42	1.06±0.29
2	0.68±0.23	2.06±1.15	1.05±0.29
3	0.75±0.26	2.49±1.35	1.05±0.30
4	1.07±0.60	2.31±1.30	1.05±0.30

Table 10.4: Experiment 2 Span Analysis using Reconstruction Precision, Leave-out Accuracy, and Estimated Accuracy

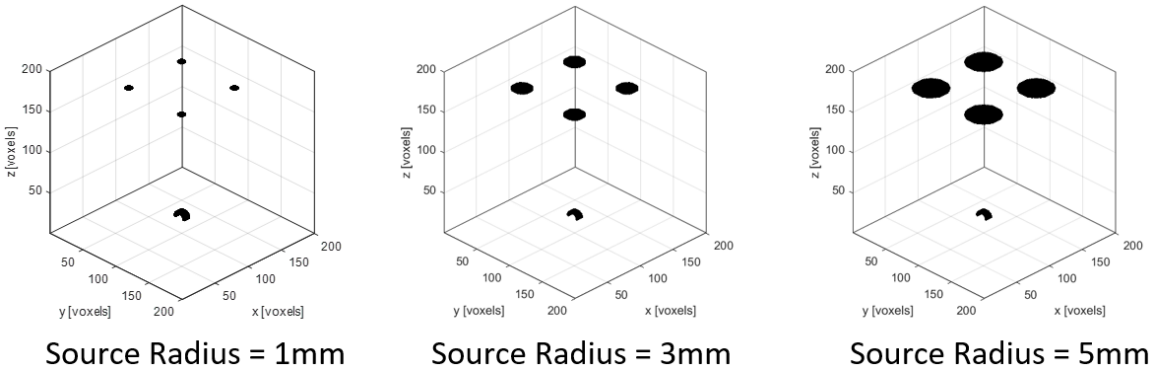


Figure 10.11: Evolution of sources and sensor in simulation environment with varying source radius

sound is fixed and only the assumed speed of sound used in algorithm is varied. The results can be seen in table 10.8.

In addition to simulating parameters related to the physical apparatus, we also want to quantify the sensitivity of this localization method to errors or uncertainties in the ToFs or distance measurements between the sensor and each of the PA markers. We simulate the scenario where we have a set of one hundred and twenty one PA sources divided in a 2cm square. The sensor is then placed at a distance of 9cm perpendicular to this square. We then add a random error selected from a zero-mean

Radius	1 (mm)	2 (mm)	3 (mm)	4 (mm)	5 (mm)
X	0	0.16	0.16	0.16	0
Y	0	0	0	0	0
Z	-0.16	-0.04	0.20	0.44	0.72
Norm	0.16	0.16	0.26	0.47	0.72

Table 10.5: Change in accuracy due to change in simulated source radius

uniform distribution to each of the distances between a PA source and the sensor. We vary the maximum absolute error. The results can be seen in table 10.9. We choose the position parameters based on our *in vivo* scenario.

10.6.3 *In Vivo*

Figure 10.14 shows a possible visualization result when using this method. A distance map is created on the surface, with the aim to give the surgeon some two-dimensional information about where the catheter tip may be relative to the surface. The three other windows show the real-time segmentation results during data collection.

Figure 10.15 shows 3 perpendicular slices of a CT volume corresponding to our computed piezoelectric element position. Each of the colored lines represent the 3

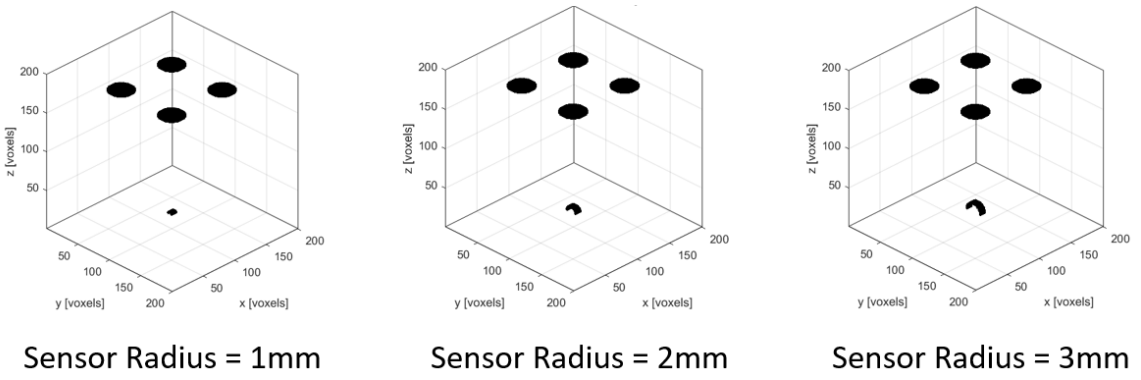


Figure 10.12: Evolution of sources and sensor in simulation environment with varying sensor radius

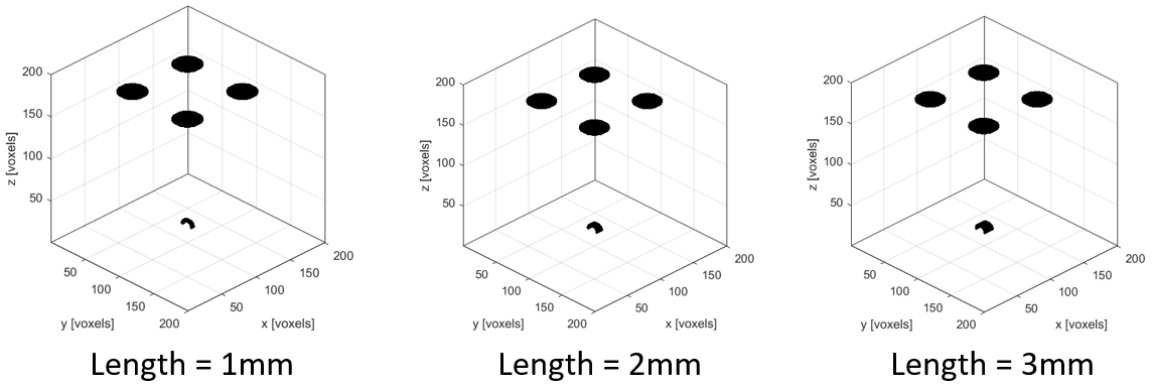


Figure 10.13: Evolution of sources and sensor in simulation environment with varying sensor length

slices (saggital, axial, coronal). Subjectively, it can be seen that the detected position is close to the position in the CT volume.

Across 10000 trials of using different subsets of 100 PA markers, the resulting metrics can be seen in Table 10.10.

Radius	1 (mm)	1.5 (mm)	2 (mm)	2.5 (mm)	3 (mm)
X	0.16	0.16	0.16	0.16	0.16
Y	0	0	0	0	0
Z	0.55	0.50	0.44	0.47	0.41
Norm	0.57	0.52	0.47	0.49	0.44

Table 10.6: Change in accuracy due to change in simulated sensor radius

10.7 Discussion

From the results of the first experiment seen in table 10.1 and 10.2, we can see that the PZT element localization results are fairly repeatable and comparable to conventional tracking-based systems. There are several sources of error that contributed to these results. Firstly, stereo camera errors are tied to how well calibrated the system is and can fluctuate depending on the specific stereo camera system used. Secondly, segmentation errors of the photoacoustic markers and acoustic signal will propagate to the final localization result. These errors are also magnified since the photoacoustic markers themselves are only several centimeters apart. Thirdly, if the photoacoustic markers are very close to each other, the quadratic system becomes less well-defined. Ideally, the markers would be spaced far apart, but there are drawbacks to doing so. The received acoustic signal by the PZT element decreases in amplitude as the

Length	1 (mm)	1.5 (mm)	2 (mm)	2.5 (mm)	3 (mm)
X	0	0	0.16	0.16	0.16
Y	0	0	0	0	0
Z	0.48	0.48	0.44	0.44	0.44
Norm	0.48	0.48	0.47	0.47	0.47

Table 10.7: Change in accuracy due to change in simulated sensor length

marker moves further away on the surface of the medium. It may become necessary to increase the energy density of the laser to counteract this effect.

Table 10.3 shows the expected trend for each of the metrics used. As the partition parameter increases, the number of points used for multilateration increases. This seems to result in an improvement in each of the metrics as each of the metrics decrease in magnitude as the partition parameter increases. However, we see that the decrease is marginal for the leave-out accuracy metric. This may indicate that there remains some segmentation noise in either or both of the stereocamera points or piezoelectric element signal.

Interestingly, table 10.4 does not show the expected trend. The only metric that shows a trend is the reconstruction precision. In this case, it is increasing as the span parameter increases. One explanation for this may be that PA spots that are

SoS	1460 (m/s)	1480 (m/s)	1500 True SoS (m/s)	1520 (m/s)	1540 (m/s)
X	0.1	0.2	0.16	0.16	0.17
Y	0	0	0	0	0
Z	1.6	1.0	0.44	-0.13	0.69
Norm	1.6	1.0	0.47	0.21	0.71

Table 10.8: Change in accuracy due to change in simulated speed of sound

Figure 10.14: Software result during *in vivo* experiment and possible visualization of a surface distance map

closer together end up resulting in a computed spot that is more consistent. We expected the metrics to improve as the span parameter increased as multilateration theoretically operates better when the PA spots are further apart. This result may again be an indication that segmentation errors eroded any effects that an increased span parameter may have had on the method.

There are a few observations to make from the simulation results. As expected,

Maximum Absolute Error (mm)	2	4	6	8	10 (mm)
Minimum	0.46	0.41	0.26	0.31	0.31
Mean	1.59	2.75	4.01	5.25	6.61
Median	1.47	2.56	3.72	4.89	6.16
Maximum	6.13	10.10	14.18	19.88	24.64
STD	0.67	1.34	2.04	2.71	3.43

Table 10.9: Simulation studying sensitivity of the localization result to errors in the source-sensor distance.

we can see from table 10.5 that the source radius has a detrimental effect on the localization result. In a sense, this causes the sources to act less like a point target, which was one of our assumptions. On the other hand, the sensor radius and sensor length has minimal effect on the resulting accuracy, as can be seen in table 10.6 and table 10.7. This may be because we apply correction to the time of flights based on our knowledge about the sensor. For example, if the sensor is 2mm in radius, 2mm is added to each of the time of flights/distances. The simulated varying speed of sound results shown in table 10.8 slightly differs from expectations in that the minimum error is not at the true SoS. We believe this is due to the partial volume effects within the simulation itself as both the source and sensor surfaces have a finite thickness associated with them. The sensitivity results shown in table 10.9 seem to indicate that this method is fairly sensitive to the maximum absolute error on the

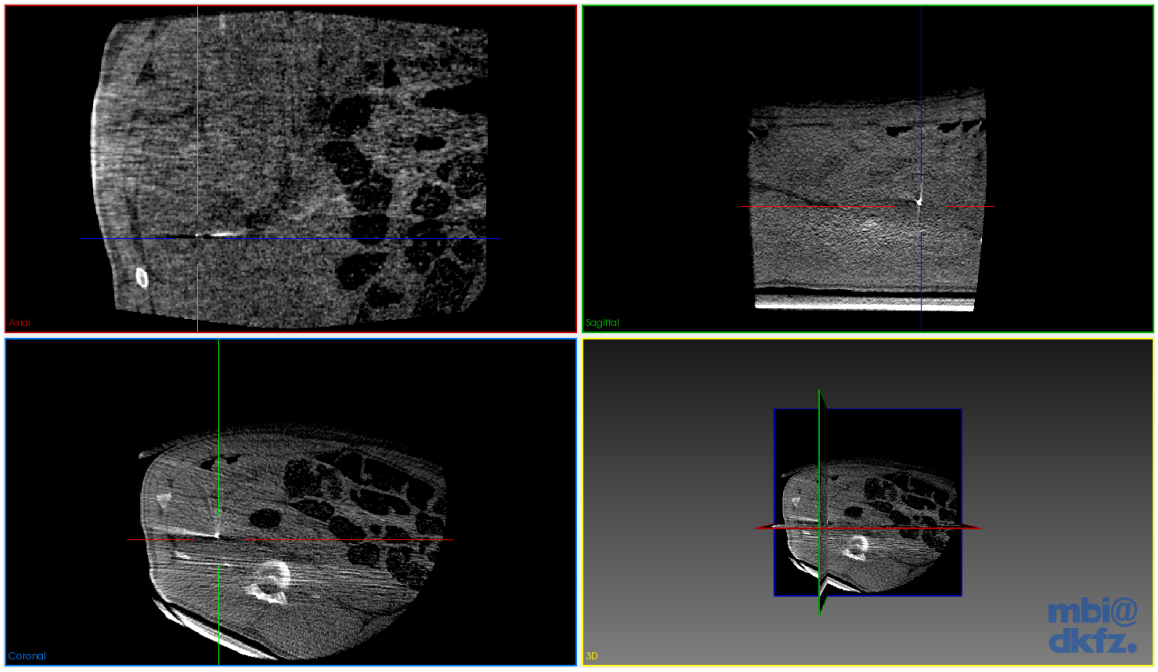


Figure 10.15: Perpendicular slices of a CT volume at the computed piezoelectric element position during *in vivo* experiment

source-sensor distances as a centimeter of error can lead to an accuracy of 6-7mm on average.

The *in vivo* experiments and results show that this method is feasible in an *in vivo* setting. While the errors are higher than the phantom experiments, they may still be useful depending on the surgical procedure. If we consider the sensitivity results in table 10.9, there may be source-sensor uncertainties of a centimeter. In addition, there are also error sources due to registration with the ground truth CT modality. Another likely source of error is tissue inhomogeneity and consequent SoS inhomogeneity. In our models, we assume a single SoS value for the ToF measurements from each PA marker. This is unlikely to be the case in a realistic setting because

Metric	RP (mm)	EA (mm)	LA (mm)
Minimum	0.05	1.73	0
Mean	2.59	8.69	1.26
Median	1.92	8.20	0.79
Maximum	12.33	18.88	18.89
STD	1.83	2.60	1.91

Table 10.10: *In vivo* experimental results

each acoustic path may differ in terms of its tissue composition. While it is impossible to solve for a unique SoS for each acoustic path, one possible future direction may be to allow for some differences in SoS for each acoustic path constrained by their relative positions to each other. Another possibility is to use pre-operative imaging as a means for estimating SoS. Certain data processing steps such as surface or speed of sound compensation, as well as improvements to the apparatus such as reducing the source radius, may reduce these errors.

Earlier, we mentioned that we sequentially collect data, one point at a time. While this is possible in a bench experiment, it is impossible in practice while the catheter

is moving. The main challenge in concurrently collecting data from multiple photoacoustic markers is distinguishing the acoustic signals in the received PZT waveform. Currently, this method requires known correspondence between the photoacoustic marker positions and its distance measure. Another direction is modeling for inhomogeneous speed of sound within the body. Further development would be necessary to improve the practicality of this method.

This localization method can have a synergistic effect on current fluoroscopic catheter tracking surgical workflows. The results shown here represent a first step towards reducing the need for continuous fluoroscopy during catheter insertion. However, before this system can be used in practice, several improvements and developments are necessary. The apparatus is currently too bulky to be used in conjunction with x-ray fluoroscopy. Reducing the size of the laser or making the light delivery system hand-held would be a necessary development. Additionally, the current data acquisition rate limits this method as a static tracking method. Improving the acquisition rate may be possible by using more sophisticated light delivery methods such as coded excitation. Increasing the tracking range by integrating a robotic light delivery manipulation component would also improve the practicality of this system.

10.8 Conclusions

We experimentally demonstrated the localization of a single piezoelectric element. The proposed system can be used to track catheters and provides photoacoustic surgical systems with an additional function. The resulting localization precision and relative accuracy are comparable with existing tracking methods. In addition, we showed a reconstruction precision up to 0.4mm and an estimated accuracy up to 0.5mm. We also presented a simulation studying the effects of the source radius, the sensor radius, the sensor length, and the speed of sound. *In vivo* results were also shown. Future work will the development of concurrent photoacoustic marker projection and its associated methods.

Chapter 11

Conclusion

In this thesis, I introduced various methods for localizing active acoustic and photoacoustic point sources for the purposes of calibration and tracking in intraoperative ultrasound. The goals of these methods awere (1) to improve localization and visualization for point targets that are not easily distinguished under conventional ultrasound and (2) to track and register ultrasound sensors with the use of active point sources as non-physical fiducials or markers.

We applied these methods to three main research topics. The first was an ultrasound calibration framework that utilizes an active acoustic source as the phantom to aid in in-plane segmentation as well as out-of-plane estimation. The second is an interventional photoacoustic surgical system that utilizes the photoacoustic effect to create markers for tracking ultrasound transducers. We demonstrated variations of this idea to track a wide range of ultrasound transducers (three-dimensional, two-

dimensional, bi-planar). The third is a set of interventional tool tracking methods combining the use of acoustic elements embedded onto the tool with the use of photoacoustic markers. These parts explored the use of active points in the context of ultrasound calibration, tracking of ultrasound transducer, and interventional tool tracking.

Appendix A

Code and Data Repositories

The data and code for the work presented in this thesis can be found on the local MUSiiC tera-station.

Bibliography

[1] G. S. Chirikjian and A. B. Kyatkin, *Engineering applications of noncommutative harmonic analysis: with emphasis on rotation and motion groups*. CRC press, 2000.

[2] O. Ukimura, C. Magi-Galluzzi, and I. S. Gill, "Real-time transrectal ultrasound guidance during laparoscopic radical prostatectomy: impact on surgical margins," *The Journal of urology*, vol. 175, no. 4, pp. 1304–1310, 2006.

[3] O. Ukimura and I. S. Gill, "Imaging-assisted endoscopic surgery: Cleveland clinic experience," *Journal of Endourology*, vol. 22, no. 4, pp. 803–810, 2008.

[4] M. Menack, J. Spitz, and M. Arregui, "Staging of pancreatic and ampullary cancers for resectability using laparoscopy with laparoscopic ultrasound," *Surgical endoscopy*, vol. 15, no. 10, pp. 1129–1134, 2001.

[5] S. W. Kwan, M. Bhargavan, R. K. Kerlan Jr, and J. H. Sunshine, "Effect of advanced imaging technology on how biopsies are done and who does them 1," *Radiology*, vol. 256, no. 3, pp. 751–758, 2010.

BIBLIOGRAPHY

[6] M. A. Choti, "Surgical management of hepatocellular carcinoma: resection and ablation," *Journal of vascular and interventional radiology*, vol. 13, no. 9, pp. S197–S203, 2002.

[7] D. R. Reed, K. E. Wallner, S. Narayanan, S. G. Sutlief, E. C. Ford, and P. S. Cho, "Intraoperative fluoroscopic dose assessment in prostate brachytherapy patients," *International Journal of Radiation Oncology* Biology* Physics*, vol. 63, no. 1, pp. 301–307, 2005.

[8] A. L. Shiloh, R. H. Savel, L. M. Paulin, and L. A. Eisen, "Ultrasound-guided catheterization of the radial artery: a systematic review and meta-analysis of randomized controlled trials," *CHEST Journal*, vol. 139, no. 3, pp. 524–529, 2011.

[9] S. D. Wexner, R. Bergamaschi, A. Lacy, J. Udo, H. Brölmann, R. H. Kennedy, and H. John, "The current status of robotic pelvic surgery: results of a multinational interdisciplinary consensus conference," *Surgical endoscopy*, vol. 23, no. 2, pp. 438–443, 2009.

[10] N. Kuo, H. J. Kang, D. Y. Song, J. U. Kang, and E. M. Boctor, "Real-time photoacoustic imaging of prostate brachytherapy seeds using a clinical ultrasound system," *Journal of biomedical optics*, vol. 17, no. 6, pp. 0 660 051–0 660 057, 2012.

[11] H. Rivaz, I. Fleming, L. Assumpcao, G. Fichtinger, U. Hamper, M. Choti,

G. Hager, and E. Boctor, "Ablation monitoring with elastography: 2d in-vivo and 3d ex-vivo studies," *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2008*, pp. 458–466, 2008.

[12] H. Rivaz, E. M. Boctor, M. A. Choti, and G. D. Hager, "Real-time regularized ultrasound elastography," *IEEE transactions on medical imaging*, vol. 30, no. 4, pp. 928–945, 2011.

[13] I. N. Fleming, H. Rivaz, K. Macura, L.-M. Su, U. Hamper, G. A. Lagoda, A. L. Burnett II, T. Lotan, R. H. Taylor, G. D. Hager *et al.*, "Ultrasound elastography: enabling technology for image guided laparoscopic prostatectomy," in *SPIE Medical Imaging*. International Society for Optics and Photonics, 2009, pp. 72 612I–72 612I.

[14] H. Rivaz, P. Foroughi, I. Fleming, R. Zellars, E. Boctor, and G. Hager, "Tracked regularized ultrasound elastography for targeting breast radiotherapy," *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009*, pp. 507–515, 2009.

[15] B. E. Kang H., Stolka PJ, "Openitglinkmusiic: A standard communications protocol for advanced ultrasound research," 2011.

[16] R. Estépar, N. Stylopoulos, R. Ellis, E. Samset, C.-F. Westin, C. Thompson, and K. Vosburgh, "Towards scarless surgery: an endoscopic-ultrasound naviga-

tion system for transgastric access procedures," *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2006*, pp. 445–453, 2006.

[17] R. S. J. Estépar, C.-F. Westin, and K. G. Vosburgh, "Towards real time 2d to 3d registration for ultrasound-guided endoscopic and laparoscopic procedures," *International journal of computer assisted radiology and surgery*, vol. 4, no. 6, p. 549, 2009.

[18] P. J. Stolka, M. Keil, G. Sakas, E. McVeigh, M. E. Allaf, R. H. Taylor, and E. M. Boctor, "A 3d-elastography-guided system for laparoscopic partial nephrectomies," in *SPIE Medical Imaging*. International Society for Optics and Photonics, 2010, pp. 76 251I–76 251I.

[19] J. Stoll and P. Dupont, "Passive markers for ultrasound tracking of surgical instruments," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2005, pp. 41–48.

[20] S. H. Okazawa, R. Ebrahimi, J. Chuang, R. N. Rohling, and S. E. Salcudean, "Methods for segmenting curved needles in ultrasound images," *Medical image analysis*, vol. 10, no. 3, pp. 330–342, 2006.

[21] R. Rohling, W. Fung, and P. Lajevardi, "Pupil: Programmable ultrasound platform and interface library," *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2003*, pp. 424–431, 2003.

[22] X. Guo, A. Cheng, H. K. Zhang, H.-J. Kang, R. Etienne-Cummings, and E. M. Boctor, "Active echo: a new paradigm for ultrasound calibration," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2014, pp. 397–404.

[23] A. Cheng, X. Guo, H. K. Zhang, H. Kang, R. Etienne-Cummings, and E. M. Boctor, "Active point out-of-plane ultrasound calibration," in *SPIE medical imaging*. International Society for Optics and Photonics, 2015, pp. 94 150W–94 150W.

[24] A. Cheng, M. K. Ackerman, G. S. Chirikjian, and E. M. Boctor, "Design and development of an ultrasound calibration phantom and system," in *Proceedings of SPIE*, vol. 9036, 2014, pp. 903 624–903 624.

[25] M. K. Ackerman, A. Cheng, and G. Chirikjian, "An information-theoretic approach to the correspondence-free ax= xb sensor calibration problem," in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE, 2014, pp. 4893–4899.

[26] M. K. Ackerman, A. Cheng, E. Boctor, and G. Chirikjian, "Online ultrasound sensor calibration using gradient descent on the euclidean group," in *Robotics and automation (ICRA), 2014 IEEE international conference on*. IEEE, 2014, pp. 4900–4905.

[27] M. K. Ackerman, A. Cheng, B. Shiffman, E. Boctor, and G. Chirikjian, "Sensor calibration with unknown correspondence: Solving ax= xb using euclidean-group

invariants," in *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*. IEEE, 2013, pp. 1308–1313.

[28] F. Aalamifar, A. Cheng, Y. Kim, X. Hu, H. K. Zhang, X. Guo, and E. M. Boctor, "Robot-assisted automatic ultrasound calibration," *International journal of computer assisted radiology and surgery*, vol. 11, no. 10, pp. 1821–1829, 2016.

[29] A. Cheng, J. U. Kang, R. H. Taylor, and E. M. Boctor, "Direct 3d ultrasound to video registration using photoacoustic effect," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, Berlin, Heidelberg, 2012, pp. 552–559.

[30] ——, "Direct three-dimensional ultrasound-to-video registration using photoacoustic markers," *Journal of biomedical optics*, vol. 18, no. 6, pp. 066 013–066 013, 2013.

[31] A. Cheng, X. Guo, H.-J. Kang, B. Tavakoli, J. U. Kang, R. H. Taylor, and E. M. Boctor, "Concurrent photoacoustic markers for direct three-dimensional ultrasound to video registration," in *Proc. of SPIE Vol*, vol. 8943, 2014, pp. 89 435J–1.

[32] A. Cheng, X. Guo, H. J. Kang, M. A. Choti, J. U. Kang, R. H. Taylor, and E. M. Boctor, "Direct ultrasound to video registration using photoacoustic markers from a single image pose," in *SPIE BiOS*. International Society for Optics and Photonics, 2015, pp. 93 130X–93 130X.

[33] A. Cheng, H. J. Kang, H. K. Zhang, R. H. Taylor, and E. M. Boctor, "Ultrasound to video registration using a bi-plane transrectal probe with photoacoustic markers," in *SPIE Medical Imaging*. International Society for Optics and Photonics, 2016, pp. 97 860J–97 860J.

[34] A. Cheng, H. K. Zhang, J. U. Kang, R. H. Taylor, and E. M. Boctor, "Localization of subsurface photoacoustic fiducials for intraoperative guidance," in *Proc. of SPIE Vol*, vol. 10054, 2017, pp. 100 541D–1.

[35] A. Cheng, Y. Itsarachaiyot, Y. Kim, H. K. Zhang, R. H. Taylor, and E. M. Boctor, "Catheter tracking in an interventional photoacoustic surgical system," in *SPIE Medical Imaging*. International Society for Optics and Photonics, 2017, pp. 1 013 527–1 013 527.

[36] A. Cheng, D. S. Lee, N. Deshmukh, and E. M. Boctor, "Virtual rigid body: a new optical tracking paradigm in image-guided interventions," in *Proc. of SPIE Vol*, vol. 9415, 2015, pp. 94 150Y–1.

[37] Q. Ma, J. D. Davis, A. Cheng, Y. Kim, G. S. Chirikjian, and E. M. Boctor, "A new robotic ultrasound system for tracking a catheter with an active piezoelectric element," in *Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on*. IEEE, 2016, pp. 2321–2328.

[38] R. W. Prager, R. N. Rohling, A. Gee, and L. Berman, "Rapid calibration for

3-d freehand ultrasound," *Ultrasound in medicine & biology*, vol. 24, no. 6, pp. 855–869, 1998.

[39] P. R. Detmer, G. Bashein, T. Hodges, K. W. Beach, E. P. Filer, D. H. Burns, and D. E. Strandness, "3d ultrasonic image feature localization based on magnetic scanhead tracking: in vitro calibration and validation," *Ultrasound in medicine & biology*, vol. 20, no. 9, pp. 923–936, 1994.

[40] N. Pagoulatos, D. R. Haynor, and Y. Kim, "A fast calibration method for 3-d tracking of ultrasound images using a spatial localizer," *Ultrasound in medicine & biology*, vol. 27, no. 9, pp. 1219–1229, 2001.

[41] E. Boctor, A. Viswanathan, M. Choti, R. H. Taylor, G. Fichtinger, and G. Hager, "A novel closed form solution for ultrasound calibration," in *Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium on.* IEEE, 2004, pp. 527–530.

[42] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE transactions on systems, man, and cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[43] D. H. Ballard, "Generalizing the hough transform to detect arbitrary shapes," *Pattern recognition*, vol. 13, no. 2, pp. 111–122, 1981.

[44] A. Myronenko and X. Song, "Point set registration: Coherent point drift," *IEEE*

transactions on pattern analysis and machine intelligence, vol. 32, no. 12, pp. 2262–2275, 2010.

[45] H. W. Kuhn, "The hungarian method for the assignment problem," *Naval Research Logistics (NRL)*, vol. 2, no. 1-2, pp. 83–97, 1955.

[46] H. H. Chen, "A screw motion approach to uniqueness analysis of head-eye geometry," in *Computer Vision and Pattern Recognition, 1991. Proceedings CVPR'91., IEEE Computer Society Conference on.* IEEE, 1991, pp. 145–151.

[47] G. S. Chirikjian and Y. Yan, "Mathematical aspects of molecular replacement. ii. geometry of motion spaces," *Acta Crystallographica Section A: Foundations of Crystallography*, vol. 68, no. 2, pp. 208–221, 2012.

[48] K. Daniilidis, "Hand-eye calibration using dual quaternions," *The International Journal of Robotics Research*, vol. 18, no. 3, pp. 286–298, 1999.

[49] N. Andreff, R. Horaud, and B. Espiau, "Robot hand-eye calibration using structure-from-motion," *The International Journal of Robotics Research*, vol. 20, no. 3, pp. 228–248, 2001.

[50] X. Guo, H.-J. Kang, R. Etienne-Cummings, and E. M. Boctor, "Active ultrasound pattern injection system (auspis) for interventional tool guidance," *PloS one*, vol. 9, no. 10, p. e104262, 2014.

BIBLIOGRAPHY

[51] P. J. Besl and N. D. McKay, "Method for registration of 3-d shapes," in *Robotics-DL tentative*. International Society for Optics and Photonics, 1992, pp. 586–606.

[52] H.-J. Kang, N. Kuo, X. Guo, D. Song, J. U. Kang, and E. M. Boctor, "Software framework of a real-time pre-beamformed rf data acquisition of an ultrasound research scanner," in *SPIE Medical Imaging*. International Society for Optics and Photonics, 2012, pp. 83 201F–83 201F.

[53] X. Guo, B. Tavakoli, H.-J. Kang, J. U. Kang, R. Etienne-Cummings, and E. M. Boctor, "Photoacoustic active ultrasound element for catheter tracking," in *SPIE BiOS*. International Society for Optics and Photonics, 2014, pp. 89 435M–89 435M.

[54] Y. Wang, S. E. Butner, and A. Darzi, "The developing market for medical robotics," *Proceedings of the IEEE*, vol. 94, no. 9, pp. 1763–1771, 2006.

[55] R. H. Taylor, S. Lavallee, G. C. Burdea, and R. Mösges, "Computer-integrated surgery: Technology and clinical applications." *Clinical orthopaedics and related research*, vol. 354, pp. 5–7, 1998.

[56] C. Cheung, C. Wedlake, J. Moore, S. Pautler, and T. Peters, "Fused video and ultrasound images for minimally invasive partial nephrectomy: a phantom study," *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010*, pp. 408–415, 2010.

[57] N. Navab, M. Mitschke, and O. Schütz, "Camera-augmented mobile c-arm (camc) application: 3d reconstruction using a low-cost mobile c-arm," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI99*. Springer, 1999, pp. 688–697.

[58] A. D. Wiles, D. G. Thompson, and D. D. Frantz, "Accuracy assessment and interpretation for optical tracking systems," in *Proceedings of SPIE*, vol. 5367, 2004, pp. 421–432.

[59] T. C. Poon and R. N. Rohling, "Comparison of calibration methods for spatial tracking of a 3-d ultrasound probe," *Ultrasound in medicine & biology*, vol. 31, no. 8, pp. 1095–1108, 2005.

[60] L. Mercier, T. Langø, F. Lindseth, and L. D. Collins, "A review of calibration techniques for freehand 3-d ultrasound systems," *Ultrasound in medicine & biology*, vol. 31, no. 2, pp. 143–165, 2005.

[61] J. Leven, D. Burschka, R. Kumar, G. Zhang, S. Blumenkranz, X. Dai, M. Awad, G. Hager, M. Marohn, M. Choti *et al.*, "Davinci canvas: a telerobotic surgical system with integrated, robot-assisted, laparoscopic ultrasound capability," *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2005*, pp. 811–818, 2005.

[62] M. C. Yip, T. K. Adebar, R. N. Rohling, S. E. Salcudean, and C. Y. Nguan, "3d ultrasound to stereoscopic camera registration through an air-tissue bound-

ary," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2010, pp. 626–634.

[63] S. Vyas, S. Su, R. Kim, N. Kuo, R. H. Taylor, J. U. Kang, and E. M. Boctor, "Intraoperative ultrasound to stereocamera registration using interventional photoacoustic imaging," in *SPIE Medical Imaging*. International Society for Optics and Photonics, 2012, pp. 83 160S–83 160S.

[64] R. G. Kolkman, W. Steenbergen, and T. G. van Leeuwen, "In vivo photoacoustic imaging of blood vessels with a pulsed laser diode," *Lasers in medical science*, vol. 21, no. 3, pp. 134–139, 2006.

[65] M. Xu and L. V. Wang, "Photoacoustic imaging in biomedicine," *Review of scientific instruments*, vol. 77, no. 4, p. 041101, 2006.

[66] C. Hoelen, F. De Mul, R. Pongers, and A. Dekker, "Three-dimensional photoacoustic imaging of blood vessels in tissue," *Optics letters*, vol. 23, no. 8, pp. 648–650, 1998.

[67] B. E. Treeby and B. T. Cox, "k-wave: Matlab toolbox for the simulation and reconstruction of photoacoustic wave fields," *Journal of biomedical optics*, vol. 15, no. 2, pp. 021 314–021 314, 2010.

[68] I. Commission *et al.*, "Safety of laser products–part 1: Equipment classification and requirements."

[69] J.-Y. Bouguet, "Camera calibration toolbox for matlab," 2004.

[70] S. Prahl, "Tabulated molar extinction coefficient for hemoglobin in water," *http://omlc. ogi. edu/spectra/hemoglobin/summary. html*, 1999.

[71] R. L. van Veen, H. Sterenborg, A. Pifferi, A. Torricelli, and R. Cubeddu, "Determination of vis-nir absorption coefficients of mammalian fat, with time-and spatially resolved diffuse reflectance and transmission spectroscopy," in *Biomedical Topical Meeting*. Optical Society of America, 2004, p. SF4.

[72] B. C. Byram, G. Holley, D. M. Giannantonio, and G. E. Trahey, "3-d phantom and in vivo cardiac speckle tracking using a matrix array and raw echo data," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 57, no. 4, 2010.

[73] M. A. L. Bell, B. C. Byram, E. J. Harris, P. M. Evans, and J. C. Bamber, "In vivo liver tracking with a high volume rate 4d ultrasound scanner and a 2d matrix array probe," *Physics in medicine and biology*, vol. 57, no. 5, p. 1359, 2012.

[74] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[75] R. O. Duda and P. E. Hart, "Use of the hough transformation to detect lines

and curves in pictures," *Communications of the ACM*, vol. 15, no. 1, pp. 11–15, 1972.

[76] J. B. West and C. R. Maurer, "Designing optically tracked instruments for image-guided surgery," *IEEE transactions on medical imaging*, vol. 23, no. 5, pp. 533–545, 2004.

[77] D. Scharstein and R. Szeliski, "High-accuracy stereo depth maps using structured light," in *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, vol. 1. IEEE, 2003, pp. I–I.

[78] J. Salvi, J. Pages, and J. Batlle, "Pattern codification strategies in structured light systems," *Pattern recognition*, vol. 37, no. 4, pp. 827–849, 2004.

[79] P. McIlroy, S. Izadi, and A. Fitzgibbon, "Kinectrack: Agile 6-dof tracking using a projected dot pattern," in *Mixed and Augmented Reality (ISMAR), 2012 IEEE International Symposium on.* IEEE, 2012, pp. 23–29.

[80] C. Wienss, I. Nikitin, G. Goebbels, K. Troche, M. Göbel, L. Nikitina, and S. Müller, "Sceptre: an infrared laser tracking system for virtual environments," in *Proceedings of the ACM symposium on Virtual reality software and technology.* ACM, 2006, pp. 45–50.

[81] J. Liu, X. Gao, Z. Zhang, S. Gao, and J. Zhou, "A new calibration method in 3d ultrasonic imaging system," in *Engineering in Medicine and Biology Society,*

1998. Proceedings of the 20th Annual International Conference of the IEEE, vol. 2. IEEE, 1998, pp. 839–841.

[82] R. H. Byrd, R. B. Schnabel, and G. A. Shultz, "Approximate solution of the trust region problem by minimization over two-dimensional subspaces," *Mathematical programming*, vol. 40, no. 1, pp. 247–263, 1988.

[83] K. Levenberg, "A method for the solution of certain non-linear problems in least squares," *Quarterly of applied mathematics*, vol. 2, no. 2, pp. 164–168, 1944.

[84] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *Journal of the society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963.

[85] K. S. Arun, T. S. Huang, and S. D. Blostein, "Least-squares fitting of two 3-d point sets," *IEEE Transactions on pattern analysis and machine intelligence*, no. 5, pp. 698–700, 1987.

[86] B. K. Horn, "Closed-form solution of absolute orientation using unit quaternions," *JOSA A*, vol. 4, no. 4, pp. 629–642, 1987.

[87] N. P. Deshmukh, H. J. Kang, S. D. Billings, R. H. Taylor, G. D. Hager, and E. M. Boctor, "Elastography using multi-stream gpu: an application to online tracked ultrasound elastography, in-vivo and the da vinci surgical system," *PloS one*, vol. 9, no. 12, p. e115881, 2014.

[88] S. Billings, N. Deshmukh, H. J. Kang, R. Taylor, and E. M. Boctor, "System for robot-assisted real-time laparoscopic ultrasound elastography," in *SPIE Medical Imaging*. International Society for Optics and Photonics, 2012, pp. 83 161W–83 161W.

[89] H. Sen, N. Deshmukh, R. Goldman, P. Kazanzides, R. H. Taylor, E. Boctor, and N. Simaan, "Enabling technologies for natural orifice transluminal endoscopic surgery (notes) using robotically guided elasticity imaging," *Medical Imaging*, p. 83161Y, 2012.

[90] F. Froghi, M. H. Sodergren, A. Darzi, and P. Paraskeva, "Single-incision laparoscopic surgery (sils) in general surgery: a review of current practice," *Surgical Laparoscopy Endoscopy & Percutaneous Techniques*, vol. 20, no. 4, pp. 191–204, 2010.

[91] Y. Ma, N. Gogin, P. Cathier, R. J. Housden, G. Gijsbers, M. Cooklin, M. O'Neill, J. Gill, C. A. Rinaldi, R. Razavi *et al.*, "Real-time x-ray fluoroscopy-based catheter detection and tracking for cardiac electrophysiology interventions," *Medical physics*, vol. 40, no. 7, 2013.

[92] B. J. Wood, H. Zhang, A. Durrani, N. Glossop, S. Ranjan, D. Lindisch, E. Levy, F. Banovac, J. Borgert, S. Krueger *et al.*, "Navigation with electromagnetic tracking for interventional radiology procedures: a feasibility study," *Journal of vascular and interventional radiology*, vol. 16, no. 4, pp. 493–505, 2005.

[93] M. A. L. Bell, N. P. Kuo, D. Y. Song, J. U. Kang, and E. M. Boctor, "In vivo visualization of prostate brachytherapy seeds with photoacoustic imaging," *Journal of biomedical optics*, vol. 19, no. 12, pp. 126 011–126 011, 2014.

Vita

Alexis Cheng was raised in Vancouver, British Columbia. He received a Bachelors degree in Electrical and Computer Engineering from the University of British Columbia in 2011, and a Masters degree in Computer Science from Johns Hopkins University in 2013. During his studies as a PhD candidate in the department of Computer Science at the Johns Hopkins University, he was involved in 7 journal articles, 19 conference publications, 8 abstracts, and 4 patents. He won the MUUSS fellowship in 2012, the best poster award at CARS in 2014, and the Professor Joel Dean Excellence in Teaching Award in 2016. His research interests include ultrasound-guided interventions, photoacoustic tracking, and surgical robotics.

Starting in August 2017, Alexis will join NIH to continue working on research related to ultrasound-guided interventions and surgical robotics.